

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 09:05:16 ; Search time 21 Seconds
(without alignments)
590.750 Million cell updates/sec

Title: US-09-726-899-3
Perfect score: 678
Sequence: 1 MSPFKYKPSLRLPETLDP.....DRKEKLIQEGKLDRTFHLVS 129

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 76.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	678	100.0	129	2 JE0383	NADH2 dehydrogenas
2	514	75.8	129	2 S28237	NADH2 dehydrogenas
3	315	46.5	133	2 JC2003	NADH ubiquinone ox
4	149	22.0	179	2 A33164	hypothetical prote
5	88.5	13.1	1078	2 S77162	DNA topoisomerase
6	82	12.1	757	2 AB0083	DNA topoisomerase
7	74	10.9	308	2 G83104	probable ferredoxi
8	74	10.9	451	2 F75177	tryptophan synthas
9	70.5	10.4	788	2 S75209	dnak protein syn
10	70	10.3	420	2 AI0921	UDP-MannNac dehydro
11	70	10.3	695	2 G64327	H+-transporting tw
12	70	10.3	1025	2 T18376	multidrug resistan
13	69.5	10.3	294	2 S22440	protein kinase (EC
14	69.5	10.3	328	2 B95404	probable ABC trans
15	69	10.2	298	2 S69523	hypothetical prote
16	69	10.2	1023	2 S12519	glutactin - fruit
17	69	10.2	1660	2 T18561	vitellogenin vit-6
18	68.5	10.1	294	2 S23095	protein kinase (EC
19	68.5	10.1	294	2 T49271	CELL DIVISION CONT
20	68.5	10.1	574	2 S57072	hypothetical prote
21	68.5	10.1	921	2 F84593	hypothetical prote
22	68.5	10.1	1330	2 A36373	hypothetical prote
23	68.5	10.1	1333	2 T38401	retrotransposabl
24	68	10.0	185	2 S12205	hypothetical prote
25	68	10.0	716	1 JC5061	macrophage-stimula
26	67.5	10.0	153	1 MNQBY	genome-linked prot
27	67.5	10.0	289	2 E75391	conserved hypothet
28	67.5	10.0	296	2 C82645	transcription regu
29	67.5	10.0	600	2 C83142	hypothetical prote

ALIGNMENTS

RESULT 1

JE0383

NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain NDUFB4 - human

N;Alternate names: NADH:ubiquinone oxidoreductase subunit NDUFB4

C;Species: Homo sapiens (man)

C;Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 03-Jun-2002

C;Accession: JE0383

R;Loeffen, J.L.C.M.; Triepels, R.H.; van den Heuvel, L.P.; Schuelke, M.; Buskens, C.A.F

Biochem. Biophys. Res. Commun. 253, 415-422, 1999

A;Title: cDNA of eight nuclear encoded subunits of NADH:ubiquinone oxidoreductase: Hume

A;Reference number: JE0379; MUID:99097250; PMID:9878551

A;Accession: JE0383

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-129 <LOE>

A;Cross-references: GB:AF044957; NID:94164445; PIDN:AAD05421.1; PID:94164446

C;Keywords: NAD; oxidoreductase

Query Match 100.0%; Score 678; DB 2; Length 129;
Best Local Similarity 100.0%; Pred. No. 8.5e-65;
Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSPFKYKPSLRLPETLDPAEYNISPETRAQAERLAIRAQLKEVLLQYNDPNRRGLI 60

Db 1 MSPFKYKPSLRLPETLDPAEYNISPETRAQAERLAIRAQLKEVLLQYNDPNRRGLI 60

Qy 61 ENPALLRWAVARTINVPNFRPTPNSIMGALCGFGPLIFIIYIIKTRDRKEKLIQEGK 120

Db 61 ENPALLRWAVARTINVPNFRPTPNSIMGALCGFGPLIFIIYIIKTRDRKEKLIQEGK 120

Qy 121 LDRTFHLSY 129

Db 121 LDRTFHLSY 129

RESULT 2

S28237

NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain B15 - bovine

C;Species: Bos primigenius taurus (cattle)

C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 03-Jun-2002

C;Accession: S28237

R;Walker, J.E.; Arizumi, J.M.; Dupuis, A.; Fearnley, I.M.; Medd, S.M.; Pil

J. Mol. Biol. 226, 1051-1072, 1992

A;Title: Sequences of 20 subunits of NADH: ubiquinone oxidoreductase from bovine heart

A;Reference number: S28237; MUID:92389317; PMID:1518044

A;Accession: S28237

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-129 <WAL>

A;Cross-references: EMBL:X64898; NID:g113; PIDN:CAA46107.1; PID:g114

C;Keywords: electron transfer; mitochondrion; NAD; oxidoreductase

Query Match 75.8%; Score 514; DB 2; Length 129;
Best Local Similarity 73.6%; Pred. No. 2.3e-47;
Matches 95; Conservative 18; Mismatches 16; Indels 0; Gaps 0;

QY 1 MSPPKYPSLRTLPETLDPAEYNI SPETRAQAERLAIRAQLKREYLLQYNDPNRGLI 60
DB 1 MSPPKYPSLRTLPETLDPAEYNI SPETRAQAERLAIRAQLKREYLLQYNDPNRGLI 60
QY 61 ENPALLERWAYARTINVPNRPRTPKNSLMGALCGFGLPIFYIYIINTERDRKEKLIQEGK 120
DB 61 EDPALVWTVYARSANIYVNPFRPNTKSLGALFGIPLGVFWYVYVFKTRDRKEKLIQEGK 120
QY 121 LDRTFHLSY 129
DB 121 LDRTFNISY 129

RESULT 3
JC2003
NADH ubiquinone oxidoreductase B15 chain like protein - chicken
N/Alternate names: murine Hox-3.1 homeobox
C/Species: Gallus gallus (Chicken)
C/Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 07-Feb-1997
C/Accession: JC2003
R/Goldberg, G.S.; Kaczmarczyk, W.
Gene 133, 233-235, 1993
A/Title: A chicken genomic DNA fragment that hybridizes to the murine Hox-3.1 homeobox
A/Reference number: JC2003; MUID:94040816; PMID:7901127
A/Accession: JC2003
A/Molecule type: DNA
A/Residues: 1-133 <GOL>
C/Genetics:
A/Introns: 67/3
A/Keywords: homeobox; phosphoprotein; sulfoprotein; transmembrane protein
F/95-112/Domain: transmembrane #status predicted <TMM>
F/30/Binding site: sulfate (Tyr) (covalent) #status predicted
F/114/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted

Query Match 46.5%; Score 315; DB 2; Length 133;
Best Local Similarity 59.0%; Pred. No. 3.5e-26;
Matches 62; Conservative 14; Mismatches 29; Indels 0; Gaps 0;

QY 5 KYPPSSLRTLPETLDPAEYNI SPETRAQAERLAIRAQLKREYLLQYNDPNRGLIENPA 64
DB 12 EYENRYVSLPAELDPATYDTFLKRAEAEIRALRKQYLLQLNTPKPRVIEDPA 71
QY 65 LLRWAYARTINVPNRPRTPKNSLMGALCGFGLPIFYIYIINTER 109
DB 72 LLRWYARTHNVPYTRPTKTSFLGAVFAIGPILFWIAAFKTER 116

RESULT 4
A33164
hypothetical protein walter - chicken
N/Alternate names: homeotic protein Hox-3.1 homolog; hypothetical 13.1K protein
C/Species: Gallus gallus (chicken)
C/Date: 28-Mar-1991 #sequence_revision 28-Mar-1991 #text_change 21-Jul-2000
C/Accession: A33164; JCI395
R/Goldberg, G.S.
submitted to the Protein Sequence Database, March 1991
A/Reference number: A33164
A/Accession: A33164
A/Status: preliminary; not compared with conceptual translation.
A/Molecule type: DNA
A/Residues: 1-179 <GOL>
A/Experimental source: strain Leghorn
R/Goldberg, G.S.; Kaczmarczyk, W.
Gene 121, 397-398, 1992
A/Title: Sequence of a novel chicken genomic DNA fragment that hybridizes to the murine
A/Reference number: JCI395; MUID:93077061; PMID:1359990
A/Accession: JCI395
A/Molecule type: DNA

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 08:59:25 ; Search time 17 Seconds
(without alignments)
356.850 Million cell updates/sec

Title: US-09-726-899-3

Perfect score: 678
Sequence: 1 MSFPKYPSSLRTLPETLDP.....DRKELIQEKLDRTHLSY 129

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	673	99.3	128	1 NBSM_HUMAN	O95168 homo sapien
2	509	75.1	128	1 NBSM_BOVIN	P48305 bos taurus
3	315	46.5	133	1 NBSM_CHICK	P48306 gallus gall
4	88.5	13.1	1078	1 GYRB_SYNY3	P77966 synchocyst
5	74	10.9	451	1 TRB2_PYRAB	O9v150 pyrococcus
6	71	10.5	270	1 PFG_ZYMOO	Q9x3x1 zymomonas m
7	70.5	10.4	771	1 DNK3_SYNY3	P73098 synchocyst
8	70	10.3	384	1 SX18_HUMAN	P35713 homo sapien
9	70	10.3	420	1 WCC3_SALTI	Q82389 salmonella
10	70	10.3	695	1 VATI_METUA	Q57675 methanococ
11	69.5	10.3	294	1 CC21_ORYSA	P29618 oryza sativ
12	69	10.2	298	1 VPO_BHPI1	P51719 bacterioph
13	69	10.2	656	1 MTHR_HUMAN	P42898 homo sapien
14	69	10.2	759	1 SC63_HUMAN	Q9ugp8 homo sapien
15	69	10.2	759	1 SC63_MOUSE	Q8vhe0 mus musculu
16	69	10.2	1023	1 GLT_DROME	P33438 drosophila
17	69	10.2	1660	1 VIT6_OSCER	Q94637 oscheilus br
18	68.5	10.1	294	1 CC2A_ARATH	P24100 arabidopsis
19	68.5	10.1	574	1 BFAL_YEAST	P47113 saccharomyc
20	68.5	10.1	1333	1 RT21_SCHPO	Q05654 schizosacch
21	68.5	10.1	1333	1 RT22_SCHPO	Q9c0r2 schizosacch
22	68.5	10.1	1333	1 RT23_SCHPO	Q9ur07 schizosacch
23	68	10.0	270	1 LPSC_RHIME	Q9r9m9 rhizobium m
24	67.5	10.0	153	1 VPG_BIDVP	P09513 barley yell
25	67	9.9	232	1 NEPI_PYRAB	Q9v0m0 pyrococcus
26	67	9.9	455	1 P2X5_RAT	P51578 rattus norv
27	67	9.9	618	1 M3K2_HUMAN	Q9v2u5 homo sapien
28	66.5	9.8	1036	1 SECA_SPIOL	Q36795 spinacia ol
29	66	9.7	193	1 LOLE_NEIMA	P57023 neisseria m
30	66	9.7	406	1 VPS_P8INU	Q96lx9 psilorum nu
31	66	9.7	420	1 WBCC_SALTY	Q916r4 salmonella
32	66	9.7	657	1 V94B_FOWPV	P17355 fowlpox vir
33	66	9.7	677	1 RN14_YEAST	P25298 saccharomyc

34 65.5 9.7 450 1 DBH4_LACBI
35 65.5 9.7 2358 1 MOKD_SCHPO
36 64.5 9.5 212 1 DSEA_BURCE
37 64.5 9.5 315 1 RSEB_HAEIN
38 64.5 9.5 654 1 MTHR_MOUSE
39 64.5 9.5 842 1 XJE7_YEAST
40 64.5 9.5 1003 1 ATC_ARTSP
41 64.5 9.5 1247 1 NOS_ANOST
42 63.5 9.4 294 1 CDC2_MAIZE
43 63.5 9.4 425 1 L756_CABEL
44 63.5 9.4 461 1 MANA_EWENI
45 63.5 9.4 686 1 KLC_STRPU

P54388 laccaria bi
Q9y19 schizosacch
Q9rv8 burkholderi
P44792 haemophilus
Q9w20 mus musculu
P47050 saccharomyc
P35316 artemia san
O61608 anopheles s
P23111 zea mays (m
Q11184 caenorhabdi
P29951 emericeila
Q05090 strongyloce

ALIGNMENTS

RESULT 1
NBSM_HUMAN
AC O95168; STANDARD; PRT; 128 AA.
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE NADH-ubiquinone oxidoreductase B15 subunit (EC 1.6.5.3) (EC 1.6.99.3)
DE (Complex I-B15) (CI-B15).
GN NDUFB4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99097250; PubMed=9878551;
RA Loeffen J.L.C.M., Triefels R.H., van den Heuvel L., Schuelke M.,
RA Buskens C.A.F., Smeets R.J.P., Trijbels J.M.F., Smeitink J.A.M.;
RT "cDNA of eight nuclear encoded subunits of NADH:ubiquinone
RT oxidoreductase: human complex I cDNA characterization completed.";
RL Biochem. Biophys. Res. Commun. 253:415-422(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Cervix;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hong L.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield A.S., Krzyzinski M.I., Skalska U., Smalish D.E.,
RA Schnerch A., Schen J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: TRANSFER OF ELECTRONS FROM NADH TO THE RESPIRATORY
CC CHAIN. THE IMMEDIATE ELECTRON ACCEPTOR FOR THE ENZYME IS BELIEVED
CC TO BE UBIQUINONE.
CC -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -!- CATALYTIC ACTIVITY: NADH + acceptor = NAD(+) + reduced acceptor.
CC -!- SUBUNIT: COMPLEX I IS COMPOSED OF ABOUT 40 DIFFERENT SUBUNITS.
CC -!- SUBCELLULAR LOCATION: Mitochondrial inner membrane; matrix side.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration -
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
CC EMBL; AF044957; AAD05421.1; --
CC DR EMBL; BC000855; AAH00855.1; --
CC DR PIR; J03083; J03083.
CC DR Genew; HGNC:7699; NDUFB4.
CC DR MIM; 603840; --
CC DR GO; GO:0005748; C:NADH dehydrogenase complex (ubiquinone) (se. .; TAS.
CC DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; TAS.
CC DR GO; GO:0006120; P:oxidative phosphorylation, NADH to ubiquinone; TAS.
CC KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion; Acetylation.
CC FT INIT MET 0
CC FT MOD_RES 1 1 ACETYLATION (BY SIMILARITY).
CC SQ SEQUENCE 128 AA; 15077 MW; DCCF6BF46F9D8 CRC64;

Query Match 99.3%; Score 673; DB 1; Length 128;
Best Local Similarity 100.0%; Pred. No. 1.3e-62;
Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SPFKYKPSRLTLPETLPAEYINISPTETRAQERLAIRAQLKREYLLQYNDPNNRGLIE 61
DB 1 SPFKYKPSRLTLPETLPAEYINISPTETRAQERLAIRAQLKREYLLQYNDPNNRGLIE 60

QY 62 NPALLRWAYARTINYPNRPFPKNSLMGALCGPGLPIFYIYIITRDRKKEKLIQEGKL 121
DB 61 NPALLRWAYARTINYPNRPFPKNSLMGALCGPGLPIFYIYIITRDRKKEKLIQEGKL 120

QY 122 DRTFHSY 129
DB 121 DRTFHSY 128

RESULT 2
NB5M_BOVIN
ID NB5M_BOVIN STANDARD; PRT; 128 AA.
AC P48305;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE NADH-ubiquinone oxidoreductase B15 subunit (EC 1.6.5.3) (EC 1.6.99.3)
DE (Complex I-B15) (CI-B15).
GN NDUFB4.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Heart;
RX MEDLINE=92389317; PubMed=1518044;
RA Walker J.E., Arizmenid J.M., Dupuis A., Fearnley I.M., Finel M.,
RA Medd S.M., Pilkington S.J., Runswick M.J., Skehel J.M.;
RT "Sequences of 20 subunits of NADH:ubiquinone oxidoreductase from
RT bovine heart mitochondria. Application of a novel strategy for
RT sequencing proteins using the polymerase chain reaction.";
RL J. Mol. Biol. 226:1051-1072(1992).
CC -!- FUNCTION: TRANSFER OF ELECTRONS FROM NADH TO THE RESPIRATORY
CC CHAIN. THE IMMEDIATE ELECTRON ACCEPTOR FOR THE ENZYME IS BELIEVED
CC TO BE UBIQUINONE.
CC -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -!- CATALYTIC ACTIVITY: NADH + acceptor = NAD(+) + reduced acceptor.
CC -!- SUBUNIT: COMPLEX I IS COMPOSED OF ABOUT 40 DIFFERENT SUBUNITS.
CC -!- SUBCELLULAR LOCATION: Mitochondrial inner membrane; matrix side.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
CC EMBL; X64898; CAA46107.1; --
CC DR PIR; S28237; S28237.
CC KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion; Acetylation.
CC FT INIT MET 0
CC FT MOD_RES 1 1 ACETYLATION.
CC SQ SEQUENCE 128 AA; 15053 MW; CCI352E9E80DF7D5 CRC64;

Query Match 75.1%; Score 509; DB 1; Length 128;
Best Local Similarity 73.4%; Pred. No. 1.1e-45;
Matches 94; Conservative 18; Mismatches 16; Indels 0; Gaps 0;

QY 2 SPFKYKPSRLTLPETLPAEYINISPTETRAQERLAIRAQLKREYLLQYNDPNNRGLIE 61
DB 1 SPFKYKPSRLTLPETLPAEYINISPTETRAQERLAIRAQLKREYLLQYNDPNNRGLIE 60

QY 62 NPALLRWAYARTINYPNRPFPKNSLMGALCGPGLPIFYIYIITRDRKKEKLIQEGKL 121
DB 61 DPALVRWTYARSANIYPNRPFPKNSLMGALCGPGLPIFYIYIITRDRKKEKLIQEGKL 120

QY 122 DRTFHSY 129
DB 121 DRTFHSY 128

RESULT 3
NB5M_CHICK
ID NB5M_CHICK STANDARD; PRT; 133 AA.
AC P48306;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative NADH-ubiquinone oxidoreductase B15 subunit (EC 1.6.5.3)
DE (EC 1.6.99.3) (Complex I-B15) (CI-B15) (Hypothetical protein Walter
DE (GGHPW).
GN NDUFB4.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94040816; PubMed=7901127;
RA Goldberg G.S., Kaczmarczyk W.;
RT "A chicken genomic DNA fragment that hybridizes to the murine Hox-3.1
RT homeobox is likely to encode the NADH ubiquinone oxidoreductase
RT subunit B15.";
RL Gene 133:233-235(1993).
RN [2]

RP SEQUENCE FROM N.A.
RX MEDLINE=93077061; PubMed=1359990;
RA Goldberg G.S., Kaczmarczyk W.;
RT "Sequence of a novel chicken genomic DNA fragment that hybridizes to
RT the murine Hox-3.1 homeobox.";
RL Gene 121:397-398(1992).
CC -!- FUNCTION: TRANSFER OF ELECTRONS FROM NADH TO THE RESPIRATORY
CC CHAIN. THE IMMEDIATE ELECTRON ACCEPTOR FOR THE ENZYME IS BELIEVED
CC TO BE UBIQUINONE (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -!- CATALYTIC ACTIVITY: NADH + acceptor = NAD(+) + reduced acceptor.
CC -!- SUBUNIT: COMPLEX I IS COMPOSED OF ABOUT 40 DIFFERENT SUBUNITS.
CC -!- SUBCELLULAR LOCATION: Mitochondrial inner membrane; matrix side
CC (By similarity).

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	569	83.9	119	4	Q8N4D3	Q8N4D3 homo sapien
2	529	78.0	129	11	Q9COC7	Q9COC7 mus musculu
3	236	34.8	119	11	Q9DBH2	Q9DBH2 mus musculu
4	115.5	17.0	172	5	Q23098	Q23098 caenorhabdi
5	96.5	14.2	113	5	Q3V753	Q3V753 drosophila
6	82	12.1	386	2	Q9L786	Q9L786 yersinia pe
7	82	12.1	757	16	Q8Z143	Q8Z143 yersinia pe
8	78	11.5	764	17	Q8T2H8	Q8T2H8 pyrococcus
9	75.5	11.1	980	16	Q982K5	Q982K5 rhizobium 1
10	74.5	11.0	313	2	Q9K138	Q9K138 bacteroides
11	74	10.9	173	16	Q8FWV0	Q8FWV0 corynebacte
12	74	10.9	195	11	Q9W1N8	Q9W1N8 mus musculu
13	74	10.9	308	16	Q9HW70	Q9HW70 pseudomonas
14	74	10.9	526	2	Q9WXH9	Q9WXH9 lactobacill
15	73	10.8	199	16	Q8D1Z8	Q8D1Z8 synechococc
16	73	10.8	634	13	Q90Y16	Q90Y16 fugu rubrip

DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE 0610006N12Rik protein.
 GN 0610006N12Rik.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Kidney, and Tongue;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishii K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Saïto T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaïdo I., Pesole G., Quackenbush J.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kogiwa H.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaïdo I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
 RA Hayashizaki Y.,
 RA "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 DR EMBL; AK018707; BAB2651.1; --
 DR EMBL; AK009807; BAB2651.1; --
 DR MGD; MGI:1915444; 0610006N12Rik.
 SQ SEQUENCE 129 AA; 15081 MW; 9E76B5087F095062 CRC64;

Query Match 78.0%; Score 529; DB 11; Length 129;
 Best Local Similarity 75.2%; Pred. NO. 8.8e-50;
 Matches 97; Conservative 16; Mismatches 16; Indels 0; Gaps 0;
 QY 1 MSFPKPKSLRTLPETLPDAEYNISPTERRAQAERLAIRAKREYLLQYNDPNERGLI 60
 DB 1 MGSKYKAPLAPLSTLPDAEYDVSPTERRAQAERLSIRAKREYLLQYNDPKRVSHI 60
 QY 61 ENPALLWAYARTINYPNFRPTPKNSLMGALCGFGLPIFYIYIKTERDRKEKLIQEGK 120
 DB 61 EDPALIRWTYARGANIPNFRPTPKNSLLGAVAGFGLPIFYIYVFKTRDRKERLIQEGK 120
 QY 121 LDRTHLSY 129.
 DB 121 LDRKFNISY 129

RESULT 3
 ID Q9DBH2 PRELIMINARY; PRT; 119 AA.
 AC Q9DBH2;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DE 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DT 1300010H20Rik protein.
 GN 1300010H20Rik.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Liver;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,

RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishii K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saïto T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kogiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaïdo I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
 RA Hayashizaki Y.,
 RA "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 DR EMBL; AK004958; BAB23699.1; --
 DR MGD; MGI:1915692; 1300010H20Rik.
 SQ SEQUENCE 119 AA; 13108 MW; 75801D6F5768ADC7 CRC64;

Query Match 34.8%; Score 236; DB 11; Length 119;
 Best Local Similarity 53.9%; Pred. NO. 6.5e-18;
 Matches 55; Conservative 9; Mismatches 22; Indels 16; Gaps 3;
 QY 1 MSFPKPKSLRTLPETLPDAEYNISPTERRAQAERLAIRAKREYLLQYNDPNERGLI 60
 DB 1 MGSKYKAPLAPLSTLPDAEYDVSPTERRAQAERLSIRAKREYLLQYNDPKRVSHI 60
 QY 61 ENPA-----LLRWAY-----RTINVYPN-FRPTPKN 86
 DB 61 VSGAGRGCGQPGPHLAHAGPPSLRVLDTGPEHPSPWPFN 102
 RESULT 4
 ID Q23098 PRELIMINARY; PRT; 172 AA.
 AC Q23098;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical protein W01A8.4.
 GN W01A8.4.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OC NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wilkinson J.
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C.elegans: A platform for
 RT investigating biology."
 RL Science 282:2012-2018(1998).
 DR EMBL; Z71267; CAA95851.2; --
 KW Hypothetical protein.
 SQ SEQUENCE 172 AA; 20373 MW; 26A5F61C594CEBC7 CRC64;

Query Match 17.0%; Score 115.5; DB 5; Length 172;
 Best Local Similarity 28.9%; Pred. NO. 0.00014;
 Matches 33; Conservative 23; Mismatches 45; Indels 13; Gaps 4;
 QY 16 ETLDPAEYNISPTERRAQAERLAIRAKREYLLQYNDPNERGLIENPALLRWAY 71
 DB 37 EFLPGHEYNLSDEKKAVALRWYRVEIKKLELRREYDPSFKYKEGVTMDPAMFRW-YS 95
 QY 72 RTINVYPNFRPTPKNSLMGALCGFGLPIFYIYI-----IKTERDRKEKLIQEGK 121

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 08:58:40 ; Search time 42 Seconds
(without alignments)
487.517 Million cell updates/sec

Title: US-09-726-899-3

Perfect score: 678
Sequence: 1 MSFPKYPSSLRTLPETLDP.....DRKEKLIQEGKLDRTFFHLSY 129

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq 19Jun03.*

```

1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	678	100.0	129	19 AAW69225	NADH dehydrogenase
2	530	78.2	294	22 ABB12440	Human bone marrow
3	522	77.0	117	23 ABP42296	Human ovarian anti
4	503	74.2	113	20 AAY76629	Human ovarian tumo
5	473	69.8	119	22 AAU33022	Novel human secre
6	408	60.2	115	22 AAU33023	Novel human secre
7	267	39.4	68	22 ABG22982	Novel human diagno
8	124	18.3	33	22 AAG73633	Human colon cancer
9	96.5	14.2	113	22 ABB67497	Drosophila melanog

10	95.5	14.1	155	22 ABB12358	Human bone marrow
11	78	11.5	764	22 AAB62028	Recombinant P. fur
12	76	11.2	16	21 AAB53504	Human colon cancer
13	74	10.9	451	22 AAB96680	Putative P. abyssal
14	72.5	10.7	1746	22 ABG05869	Novel human diagno
15	72	10.6	290	21 AAG33365	Zea mays protein f
16	70.5	10.4	936	23 ABP73605	Candida albicans e
17	70	10.3	432	23 ABG95649	Human nucleic acid
18	69.5	10.3	1025	23 ABP52133	Plasmidom falcipa
19	69.5	10.3	120	22 AAO07882	Human polyptide
20	69.5	10.3	292	21 AAB35797	protein involved i
21	69.5	10.3	294	20 AAW95690	Oryza sativa Cdc2
22	69	10.2	656	17 AAR88358	Human methylene-te
23	69	10.2	656	22 AAB12607	Human methylene-te
24	69	10.2	656	23 AAO15894	Human 5,10-methyle
25	69	10.2	660	17 AAR88359	Human methylene-te
26	69	10.2	660	21 AAY96186	Human methylene-te
27	69	10.2	660	22 AAE12606	Human methylene-te
28	69	10.2	660	23 AAU75413	Human methylene-te
29	69	10.2	660	23 AAU75421	Human methylene-te
30	69	10.2	660	23 AAU75422	Human methylene-te
31	69	10.2	660	23 AAU75424	Human methylene-te
32	69	10.2	660	23 AAU75425	Human methylene-te
33	69	10.2	660	23 AAU75426	Human methylene-te
34	69	10.2	660	23 AAU75428	Human methylene-te
35	69	10.2	660	23 AAU75429	Human methylene-te
36	69	10.2	660	23 AAU75430	Human methylene-te
37	69	10.2	760	22 AAW93803	Human polyptide,
38	69	10.2	760	22 AAG66537	Human interferon-a
39	69	10.2	1026	22 ABB64339	Drosophila melanog
40	68.5	10.1	204	21 AAG54044	Arabidopsis thalia
41	68.5	10.1	237	21 AAG11215	Arabidopsis thalia
42	68.5	10.1	237	21 AAG54043	Arabidopsis thalia
43	68.5	10.1	294	21 AAG11214	Arabidopsis thalia
44	68.5	10.1	294	21 AAG54042	Arabidopsis thalia
45	68.5	10.1	350	21 AAG11213	Arabidopsis thalia

ALIGNMENTS

RESULT 1

AAW69225

ID AAW69225 standard; Protein; 129 AA.

XX

AC AAW69225;

XX

DT 18-FEB-1999 (first entry)

XX

DE NADH dehydrogenase subunit NDS-2.

XX

KW NADH dehydrogenase subunit; sympathetic nervous system disease; myopathy;
KW nicotinamide-adenine dinucleotide dehydrogenase; cancer; leukaemia;
KW immune system disorder; neurodegenerative disease; therapy; NDS-2.

XX Homo sapiens.

XX

FN WO9831815-A2.

XX

PD 23-JUL-1998.

XX

PF 17-DEC-1997; 97WO-US23970.

XX

PR 17-JAN-1997; 97US-0785065.

XX (INCY-) INCYTE PHARM INC.

XX

PI Bandman O, Goli SK, Hillman JL;

XX

DR WPI; 1998-414112/35.

XX

DR N-PSDB; AAV44787.

XX

PT Human nicotinamide-adenine di-nucleotide dehydrogenase sub:units -

PD 03-JAN-2002.
 XX
 PF 07-JUN-2001; 2001WO-US18569.
 XX
 PR 07-JUN-2000; 2000US-209467P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Birse CE, Rosen CA;
 XX
 DR WPI; 2002-147878/19.
 DR N-PSDB; ABQ55373.
 XX
 PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,
 PT useful in the prevention, treatment and diagnosis of cancer (e.g.
 PT ovarian cancer), immune disorders, cardiovascular disorders and
 PT neurological diseases -
 XX
 PS Claim 11; SEQ ID No 3428; 2922pp; English.
 CC
 CC The invention relates to 2175 novel human ovarian antigens (ABP41054-
 CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical
 CC to the sequences of the invention. The invention additionally relates to
 CC recombinant vectors and host cells comprising human ovarian antigen
 CC polynucleotides, antibodies against human ovarian antigens, and the use
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
 CC treating, prognosing or preventing various ovary and/or breast-related
 CC disorders. Such conditions include ovarian cancer and breast cancer, and
 CC metastatic tumours of ovarian or breast origin, reproductive system
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
 CC vaginitis), immune disorders (e.g., congenital and acquired
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
 CC respiratory disorders, neurological disorders, gastrointestinal disorders
 CC and urinary system disorders. Ovarian antigen polypeptides and
 CC polynucleotides may also be used in screening for compounds which
 CC modulate ovarian antigen expression or activity. The polynucleotides may
 CC further be used for gene therapy, chromosome mapping, in the
 CC identification of individuals and in forensic analysis, and the
 CC polypeptides may be used as food additives or to prepare antibodies
 CC useful in disease diagnosis, drug targeting and phenotyping. The present
 CC sequence represents a human ovarian antigen of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 117 AA;
 Query Match 77.0%; Score 522; DB 23; Length 117;
 Best Local Similarity 99.0%; Pred. No. 1.1e-54;
 Matches 99; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 SPFKYKSSLTLPETLDPABYNISPTTRQAERLAIRAQKREYLLQYNDPNRRGLIE 61
 DB 18 SPFKYKSSKTLTLPETLDPABYNISPTTRQAERLAIRAQKREYLLQYNDPNRRGLIE 77
 QY 62 NPALLRWAYARTINNVNFRPTPKNSLMGALCGFGPLIFI 101
 DB 78 NPALLRWAYARTINNVNFRPTPKNSLMGALCGFGPLIFI 117
 RESULT 4
 AAY76629
 ID AAY76629 standard; Protein; 113 AA.
 XX
 AC AAY76629;
 XX
 DT 10-APR-2000 (first entry)
 XX
 XX

DE Human ovarian tumor EST fragment encoded protein 125.
 XX
 KW Expressed sequence tag; EST; human; ovarian tumor; anticancer;
 XX gene therapy; treatment.
 OS Homo sapiens.
 XX
 FN DE19817557-A1.
 XX
 PD 21-OCT-1999.
 XX
 PF 09-APR-1998; 98DE-1017557.
 XX
 PR 09-APR-1998; 98DE-1017557.
 XX
 PA (META-) METAGEN GES GENOMFORSCHUNG MBH.
 XX
 PI Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;
 XX
 DR WPI; 1999-591920/51.
 DR N-PSDB; AAZ77502.
 XX
 PT New nucleic acid sequences expressed in ovarian, and some other, cancer
 PT tissues, and derived polypeptides, for treatment of ovarian cancer and
 PT identification of therapeutic agents -
 XX
 PS Claim 25; Page 295; 310pp; German.
 CC
 CC This invention describes novel nucleic acid (cDNA) sequences (A) which
 CC have anticancer activity and are highly expressed in ovarian tumor
 CC tissue (and some also in testis and breast cancer tissue). The products
 CC of the invention can be used for gene therapy. (A) are used (i) for
 CC recombinant expression of polypeptides (B) and (ii) to isolate complete
 CC genes. (B) are used (i) to identify agents suitable for treatment of
 CC ovarian cancer; (ii) directly for treating this form of cancer
 CC (including expression from gene therapy vectors) and (iii) for generation
 CC of specific antibodies. (A) are identified by assembling ESTs (expressed
 CC sequence tags) from a particular tissue type before comparison of
 CC expression patterns. This allows a significantly longer fragment of the
 CC gene to be revealed, so should reduce the number of failures associated
 CC with the fact that ESTs from different libraries may represent different
 CC parts of the same unknown gene, distorting the estimated frequency of
 CC occurrence in a particular tissue. AAY76505-Y76638 represent protein
 CC fragments encoded by the human ovarian tumor cDNA library derived EST
 CC fragments represented in AAZ77450-Z77572.
 XX
 SQ Sequence 113 AA;
 Query Match 74.2%; Score 503; DB 20; Length 113;
 Best Local Similarity 99.0%; Pred. No. 2e-52;
 Matches 95; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 34 AERLAIRAQKREYLLQYNDPNRRGLIENPALLRWAYARTINNVNFRPTPKNSLMGALC 93
 DB 18 SERLAIRAQKREYLLQYNDPNRRGLIENPALLRWAYARTINNVNFRPTPKNSLMGALC 77
 QY 94 GFGPLIFIIYIIKTERDRKKEKLIQEGKLDRTFHLISY 129
 DB 78 GFGPLIFIIYIIKTERDRKKEKLIQEGKLDRTFHLISY 113
 RESULT 5
 AAU33022
 ID AAU33022 standard; Protein; 119 AA.
 XX
 AC AAU33022;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Novel human secreted protein #3513.
 XX
 KW Human; vaccination; gene therapy; nutritional supplement;
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;

KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
 XX Homo sapiens.
 OS WO200179449-A2.
 PN 25-OCT-2001.
 XX 16-APR-2001; 2001WO-US08656.
 XX 18-APR-2000; 2000US-0552929.
 XX 26-JAN-2001; 2001US-0770160.
 PD (HYSE-) HYSEQ INC.
 PF Tang YT, Liu C, Drmanac RT;
 XX WPI; 2001-611725/70.
 XX Nucleic acids encoding a range of human polypeptides, useful in genetic
 PI vaccination, testing and therapy -
 XX Claim 20; Page 699; 765pp; English.
 XX The invention relates to novel human secreted polypeptides. The
 CC polypeptides and antibodies to the polypeptides are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of polypeptide. The polypeptides are also useful for
 CC identifying agents (agonists and antagonists) that bind to them. Cells
 CC expressing the proteins are useful for identifying a therapeutic agent
 CC for use in treatment of a pathology related to aberrant expression or
 CC physiological interactions of the polypeptide. Vectors comprising
 CC the nucleic acids encoding the polypeptides and cells genetically
 CC engineered to express them are also useful for producing the proteins.
 CC The proteins are useful in genetic vaccination, testing and
 CC therapy, and can be used as nutritional supplements. They may be used to
 CC increase stem cell proliferation; to regulate haematopoiesis; and in
 CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
 CC immune suppression and/or stimulation; as anti-inflammatory agents; and
 CC in treatment of leukaemias. AAU29510-AAU3304 represent the amino acid
 CC sequences of novel human secreted proteins of the invention.
 XX Sequence 119 AA;
 SQ Query Match 69.8%; Score 473; DB 22; Length 119;
 Best Local Similarity 73.9%; Pred. No. 8.9e-49;
 Matches 88; Conservative 14; Mismatches 17; Indels 0; Gaps 0;
 QY 5 KYKPSRLTLPETLDPAEYNISPETTRAAQLRAQLKREYLLQYNDPNRRGLIENPA 64
 Db 1 KYTLRLATVPPTLNPAEYNISPETTRAAQLRAQLKREYLLQYNNPNRRGLIEDPA 60
 QY 65 LLRWARTINVPNRPPTKNSLMGALCGFGLPIFYIYIINTERDKKELIOEGKLD 123
 Db 61 LIRWTYARSANVYPNRPPTKNSLLGALCAFGPLFFWYCVFKTMDRQKLIREGKLDQ 119
 RESULT 6
 ID AAU33023 standard; Protein; 115 AA.
 XX AC AAU33023;
 XX DT 18-DEC-2001 (first entry)
 XX DE Novel human secreted protein #3514.
 XX Human; vaccination; gene therapy; nutritional supplement;
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
 OS Homo sapiens.
 XX

PN WO200179449-A2.
 XX 25-OCT-2001.
 XX 16-APR-2001; 2001WO-US08656.
 XX 18-APR-2000; 2000US-0552929.
 XX 26-JAN-2001; 2001US-0770160.
 PD (HYSE-) HYSEQ INC.
 PF Tang YT, Liu C, Drmanac RT;
 XX WPI; 2001-611725/70.
 XX Nucleic acids encoding a range of human polypeptides, useful in genetic
 PI vaccination, testing and therapy -
 XX Claim 20; Page 699; 765pp; English.
 XX The invention relates to novel human secreted polypeptides. The
 CC polypeptides and antibodies to the polypeptides are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of polypeptide. The polypeptides are also useful for
 CC identifying agents (agonists and antagonists) that bind to them. Cells
 CC expressing the proteins are useful for identifying a therapeutic agent
 CC for use in treatment of a pathology related to aberrant expression or
 CC physiological interactions of the polypeptide. Vectors comprising
 CC the nucleic acids encoding the polypeptides and cells genetically
 CC engineered to express them are also useful for producing the proteins.
 CC The proteins are useful in genetic vaccination, testing and
 CC therapy, and can be used as nutritional supplements. They may be used to
 CC increase stem cell proliferation; to regulate haematopoiesis; and in
 CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
 CC immune suppression and/or stimulation; as anti-inflammatory agents; and
 CC in treatment of leukaemias. AAU29510-AAU3304 represent the amino acid
 CC sequences of novel human secreted proteins of the invention.
 XX Sequence 115 AA;
 SQ Query Match 60.2%; Score 408; DB 22; Length 115;
 Best Local Similarity 66.1%; Pred. No. 5.7e-41;
 Matches 76; Conservative 16; Mismatches 23; Indels 0; Gaps 0;
 QY 9 SSLRLTLPETLDPAEYNISPETTRAAQLRAQLKREYLLQYNDPNRRGLIENPALLRW 68
 Db 1 SSLATVPANLNPATVSGISPDNRRAAQVEQLAIRAGLKKXXYLLQYNNTRIGIIKDPALIRW 60
 QY 69 AYARTINVPNRPPTKNSLMGALCGFGLPIFYIYIINTERDKKELIOEGKLD 123
 Db 61 TVARSANVYPNRPPTKNSLLGLDLCAYGLFFWYCVFKPMDRKTLPKEGLDQ 115
 RESULT 7
 ID ABG22982 standard; Protein; 68 AA.
 XX AC ABG22982;
 XX DT 18-FEB-2002 (first entry)
 XX DE Novel human diagnostic protein #22973.
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 OS Homo sapiens.
 XX WO200175067-A2.
 XX 11-OCT-2001.
 XX 30-MAR-2001; 2001WO-US08631.

```
XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX (HYSB-) HYSEQ INC.
XX Dmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
DR N-PSDB; AAS87169.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX Claim 20; SEQ ID NO 53341; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. ABG00010-ABG30377 represent novel human
XX diagnostic amino acid sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 68 AA;
Query Match 39.4%; Score 267; DB 22; Length 68;
Best Local Similarity 100.0%; Pred. No. 2.7e-24;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 61 ENPALLRWARTINVPNFRPTKNSLMGALCGFGLFIYYIKTER 109
DB 9 ENPALLRWARTINVPNFRPTKNSLMGALCGFGLFIYYIKTER 57
RESULT 8
AAG73633
ID AAG73633 standard; Protein; 33 AA.
XX AC AAG73633;
XX 03-SEP-2001 (first entry)
XX Human colon cancer antigen protein SEQ ID NO:4397.
XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
XX colorectal carcinoma.
XX Homo sapiens.
XX WO200122920-A2.
XX 05-APR-2001.
XX 28-SEP-2000; 2000WO-US26524.
XX 29-SEP-1999; 99US-0157137.
PR 03-NOV-1999; 99US-0163280.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Ruben SM, Barash SC, Birse CE, Rosen CA;
XX WPI; 2001-235357/24.
DR N-PSDB; AAH33064.
XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
XX useful for preventing, diagnosing and/or treating colorectal cancers -
XX Claim 11; Page 6222-6223; 9803pp; English.
XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
XX cancer-associated nucleic acid molecules (N) and proteins (P), where
XX the proteins are collectively known as colon cancer antigens. The colon
XX cancer antigens have cytostatic activity and can be used in gene
XX therapy and vaccine production. N and P may be used in the prevention,
XX diagnosis and treatment of diseases associated with inappropriate P
XX expression. For example, N and P may be used to treat disorders
XX associated with decreased expression by rectifying mutations or deletions
XX in a patient's genome that affect the activity of P by expressing
XX inactive proteins or to supplement the patients own production of P.
XX Additionally, N may be used to produce the colon cancer-associated Ps,
XX by inserting the nucleic acids into a host cell and culturing the cell
XX to express the proteins. N and P can be used in the prevention, diagnosis
XX and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
XX and AAG77789 represent sequences used in the exemplification of the
XX present invention.
XX N.B. Pages 666 to 682 and page 7053 of the sequence listing were
XX missing at time of publication, meaning no sequences are present for
XX SEQ ID NO:1027 to 1052, 7921 and 7922.
XX Sequence 33 AA;
Query Match 18.3%; Score 124; DB 22; Length 33;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 106 KTERDRKEKLIQEGKLDRTPHLSY 129
DB 10 KTERDRKEKLIQEGKLDRTPHLSY 33
RESULT 9
ABB67497
ID ABB67497 standard; Protein; 113 AA.
XX AC ABB67497;
XX 26-MAR-2002 (first entry)
XX Drosophila melanogaster polypeptide SEQ ID NO 29283.
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical.
XX Drosophila melanogaster.
XX WO200171042-A2.
XX 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US09231.
XX 23-MAR-2000; 2000US-191637P.
XX 11-JUL-2000; 2000US-0614150.
XX (PEKE ) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;
```


CC maintenance protein) and at least one of the archaeal polypeptides (PCNA,
CC RRC-P38 or -55, RPA, CUC6, FEN-1, dddpase, ligase, helicase dna2, or
CC helicases 2-8). (A) And similar compositions containing different
CC combinations of accessory proteins, are used to improve performance of
CC synthesis, amplification, mutagenesis, labeling and detecting reactions,
CC e.g. for gene characterization, cloning, detection of allelic variants,
CC diagnosis and screening for disease, particularly where done by
CC polymerase chain reaction (PCR). Some of the proteins also stabilize
CC duplexes during polymerase reactions or improve exonuclease reactions,
CC for example RPA also improves specificity of nucleic acid/protein
CC interaction and PCNA improves polymerase-mediated repair processes and
CC hybridization reactions. Nucleic acids encoding the archaeal polypeptides
CC are used for recombinant production of proteins, and fragments of the
CC nucleic acid as probes and primers for screening related sequences. The
CC accessory proteins increase accuracy and efficiency of polymerase
CC reactions, allow use of lower denaturation and extension temperatures
CC (possibly isothermal processing), and improve synthesis of long targets.
CC The present sequence represents a P. furiosus recombinant helicase 5.

```

SQ      Sequence       764 AA;
Query Match           11.5%; Score 78; DB 22; Length 764;
Best Local Similarity 25.8%; Pred. No. 3.8;
Matches 33; Conservative 21; Mismatches 30; Indels 44; Gaps 7;

QY      16 ETLDPAYNISPTTAAQRALATRAQLRELYLOYNDPNRRGLIENPALLRWAYARTI- 74
          :|||:|||::|||::|||:|||||:|||||:
Db      578 KTLDVDGYIIIS-----EDVAIERKSANDFIQSIIIDGR---LFDQVKRLKEAYSRPIM 626

QY      75 -----NVYPNFRPTPKNSLMGALCG----FG-PLIP-----IYYVIKTER 109
          |||:|||::|||:|||||:|||||:|||||:
Db      627 IVEGSLSYGINRVHP-----NAIRGAIAAVTVDFGVPIIFSSPETTAQVIFIARREQ 679

QY      110 DRKEKLQ 117
          :|||::|||:
Db      680 EEREKPVR 687
```

RESULT 12	
AAB53504	
ID	AAB53504 standard; Protein; 16 AA.
XX	
XX	
AC	AAB53504;
XX	
XX	
DT	09-MAR-2001 (first entry)
XX	
XX	Human colon cancer antigen protein sequence SEQ ID NO:1044.
XX	
KW	Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW	identification; cytostatic; cardioactive; neuroprotective; vulnery;
KW	immunomodulatory; muscular; gynecological; gastrointestinal;
KW	nephrotropic; antinfective; antibacterial; gene therapy; wound;
KW	neural disorder; immune system disorder; muscular disorder;
KW	reproductive disorder; gastrointestinal disorder; renal disorder;
KW	infectious disease; cardiovascular disorder.

XX	Homo sapiens.
OS	
XX	WO200055351-A1.
XX	
XX	21-SEP-2000.
XX	
XX	08-MAR-2000; 2000WO-US05883.
PF	
XX	12-MAR-1999; 99US-0124270.
XX	(HUMA-) HUMAN GENOME SCI INC.
PA	
XX	Rosen CA, Ruben SM;
PI	
XX	WPI; 2000-587534/55.
DR	N-PSDB; AAC98261.
DR	
XX	

Colon cancer associated gene sequences, referred to as colon cancer antigens, useful for the treatment, prevention, and diagnosis of colon disorders such as colon cancer -

Claim 11; Page 1624; 2104pp; English.

AAC97991 to AAC98763 encode the human colon cancer associated proteins, called human colon cancer antigens, given in AAC95234 to AAC954006. The human colon cancer antigen can have cytostatic, cardioactive, muscular, neuroprotective, immunomodulatory, gynaecological, gastrointestinal, vulvarey, nephrotropic, antiinfective and antibacterial activities, and can be used in gene therapy. The colon cancer antigen polynucleotides, proteins and antibodies to the proteins are useful for the prevention, treatment and diagnosis of colon disorders, such as colon cancer. The polynucleotides may be used in diagnostics and research, such as for chromosome identification, and as hybridisation probes. The proteins may also be used to prevent diseases such as neural disorders, immune system disorders, muscular disorders, reproductive disorders, gastrointestinal disorders, wounds, renal disorders, infectious diseases, and cardiovascular disorders. AAC98764 to AAC98772 and AAC954007 represent sequences used in the exemplification of the present invention.

SQ Sequence 16 AA;

Query Match 11.2%; Score 76; DB 21; Length 16;
Best Local Similarity 93.8%; Pred. No. 0.039;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 13
AAB96680
ID AAB96680 standard: Protein: 451 AA.

XX	AAB96680;	
XX	AC	
XX	XX	
DT	29-OCT-2001	(first entry)
XX	XX	
DE	Putative P. abyssi	tryptophan synthase beta chain #1.
XX	XX	
KW	Hyperthermophilic	archaeon; hyperthermophilic protein.
XX	XX	
OS	Pyrococcus	abyssi.
XX	XX	
PN	FR2792651-A1.	
XX	XX	
PD	27-OCT-2000.	
XX	XX	
PF	21-APR-1999;	99FR-0005034.
XX	XX	
PR	21-APR-1999;	99FR-0005034.
XX	XX	
PA	(CNRS) CNRS CENT NAT RECH SCI.	
PA	(IFRE-) IFREMER INST FR RECH EXPL MER.	
XX	XX	
PI	Forterre P, Thierry JC, Prieur D, Dietrich J, Lecompte O;	
PI	Querellou J, Weissenbach J, Saurin W, Heilig R;	
XX	XX	
DR	WPI: 2001-126236/14.	

XX New nucleotide sequences isolated from *Pyrococcus abyssi* encode
PT proteins useful in industry -
XX
XX Claim 7; Pages 1437-1439; 1657pp; French.
PS
XX The present invention relates to the genomic sequence of *Pyrococcus*
CC *abyssi* (see AA896431 and AA941223-7) and *P. abyssi* proteins. *P. abyssi* is
CC a hyperthermophilic archaean, which is isolated from deep-sea
CC hydrothermal vents. The present sequence is one such *P. abyssi* protein.
CC


```

PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match      10.6%; Score 72; DB 21; Length 290;
Best Local Similarity 26.0%; Pred. No. 5,6;
Matches 32; Conservative 15; Mismatches 44; Indels 32; Gaps 6;

Qy 4 PKYKPSLRTLPETDPAEYNISPETRAQAERLAIRAQLKREYLLQYN--DPNRR----- 57
Db 10 PLFCFGPLVARP-----PA--SSSPAGRPASQTLARRAREKLDAMFGYRKADPDLEAGGS 63
Qy 58 -----GLIENPALLWYARTINVPNFR-----PTPKNSLMGALCGFGPL 98
Db 64 SLLYPGMTESPE-LRWAFVRKIYVILTVOAMTAASAFVVKVPAVSNNFFVSSNAGIALY 122
Qy 99 IFI 101
Db 123 IFL 125

```

Search completed: January 29, 2004, 09:07:58
Job time : 43 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 09:08:31 ; Search time 34 Seconds

(without alignments)
788.509 Million cell updates/sec

Title: US-09-726-899-3

Perfect score: 678

Sequence: 1 MSPFKYKPSRLTLPETLDP.....DRKELIQEKLDRTHLSY 129

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 789580 seqs, 207824079 residues

Total number of hits satisfying chosen parameters: 789580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications AA:
- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
 - 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
 - 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	678	100.0	129	9	US-09-726-899-3
2	522	77.0	117	12	Sequence 3, Appli
3	514	75.8	129	9	Sequence 3428, Ap
4	124	18.3	33	15	Sequence 10, Appl
5	76	11.2	16	9	Sequence 4407, Ap
6	76	11.2	16	11	Sequence 1044, Ap
7	71	10.5	433	12	Sequence 1044, Ap
8	70.5	10.4	936	12	Sequence 21273, A
9	70.5	10.4	936	12	Sequence 2648, Ap
10	70	10.3	337	10	Sequence 7442, Ap
11	70	10.3	340	10	Sequence 126, App
12	70	10.3	384	10	Sequence 15, Appl
13	70	10.3	384	10	Sequence 18, Appl
14	70	10.3	370	10	Sequence 20, Appl
15	70	10.3	680	12	Sequence 100, App
					Sequence 3446, Ap

16	70	10.3	695	12	US-10-369-493-21448
17	69.5	10.3	292	12	US-10-409-701-7
18	69	10.2	656	9	US-09-728-910-4
19	69	10.2	656	10	US-09-931-795-4
20	69	10.2	656	11	US-09-802-640-34
21	69	10.2	656	12	US-10-235-172-4
22	69	10.2	656	12	US-10-316-763A-4
23	69	10.2	656	12	US-10-270-714-2
24	69	10.2	660	9	US-09-728-910-2
25	69	10.2	660	10	US-09-931-795-2
26	69	10.2	660	12	US-10-235-172-2
27	69	10.2	660	12	US-10-316-763A-2
28	69	10.2	691	12	US-10-334-143-26
29	68.5	10.1	518	12	US-10-104-047-3059
30	68.5	10.1	1214	12	US-10-369-493-2153
31	68.5	10.1	1333	12	US-10-369-493-2149
32	68.5	10.1	1333	12	US-10-369-493-2295
33	68.5	10.1	1333	12	US-10-369-493-2296
34	68.5	10.1	1333	12	US-10-369-493-2456
35	68.5	10.1	1333	12	US-10-369-493-2467
36	68.5	10.1	1333	12	US-10-369-493-22615
37	68.5	10.1	1333	12	US-10-369-493-22676
38	68	10.0	196	11	US-09-983-802-393
39	68	10.0	228	11	US-09-983-802-401
40	68	10.0	344	12	US-10-284-049-3093
41	68	10.0	365	12	US-10-287-218-12
42	68	10.0	371	11	US-09-975-719-295
43	68	10.0	724	14	US-10-068-059-12
44	68	10.0	864	15	US-10-091-333-11
45	68	10.0	864	15	US-10-325-878-11

ALIGNMENTS

RESULT 1
US-09-726-899-3
; Sequence 3, Application US/09726899
; Patent No. US20010041356A1
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goli, Surya K.
; APPLICANT: Hillman, Jennifer L.
; TITLE OF INVENTION: NOVEL SUBUNITS OF NADH DEHYDROGENASE
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESS: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Fast-SEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/726,899
; FILING DATE:
; PRIORITY DATA:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/785,065
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PP-0187 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:

Sequence 21448, A
Sequence 7, Appli
Sequence 4, Appli
Sequence 4, Appli
Sequence 34, Appli
Sequence 4, Appli
Sequence 4, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 26, Appli
Sequence 3059, Ap
Sequence 2153, Ap
Sequence 2149, Ap
Sequence 2295, Ap
Sequence 2296, Ap
Sequence 2456, Ap
Sequence 22615, A
Sequence 22676, A
Sequence 393, App
Sequence 401, App
Sequence 3093, Ap
Sequence 12, Appli
Sequence 295, App
Sequence 12, Appli
Sequence 11, Appli
Sequence 11, Appli

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 129 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: Consensus
; CLONE: Consensus
US-09-726-899-3

Query Match 100.0%; Score 678; DB 9; Length 129;
Best Local Similarity 100.0%; Pred. No. 9.4e-73; Indels 0; Gaps 0;
Matches 129; Conservative 0; Mismatches 0;

QY 1 MSFPKYPSSRLTLPETLDPAEYNISPETRAQAERLAIRAQLKREYLLQYNDPNRRGLI 60
Db 1 MSFPKYPSSRLTLPETLDPAEYNISPETRAQAERLAIRAQLKREYLLQYNDPNRRGLI 60

QY 61 ENPALLRWAYARTINYPNFRPTPKNSLMGALCGFGLPIFYIYIINTERDRKELIQEGK 120
Db 61 ENPALLRWAYARTINYPNFRPTPKNSLMGALCGFGLPIFYIYIINTERDRKELIQEGK 120

QY 121 LDRTHLSY 129
Db 121 LDRTHLSY 129

RESULT 2
US-10-264-049-3428
; Sequence 3428, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133P1
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: Patent In Ver. 3.1
; SEQ ID NO 3428
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (3)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (27)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
US-10-264-049-3428

Query Match 77.0%; Score 522; DB 12; Length 117;
Best Local Similarity 99.0%; Pred. No. 3.2e-54;
Matches 99; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 SPFKYKPSRLTLPETLDPAEYNISPETRAQAERLAIRAQLKREYLLQYNDPNRRGLIE 61
Db 18 SPFKYKPSRLTLPETLDPAEYNISPETRAQAERLAIRAQLKREYLLQYNDPNRRGLIE 77

QY 62 NPALLRWAYARTINYPNFRPTPKNSLMGALCGFGLPIFYI 101
Db 78 NPALLRWAYARTINYPNFRPTPKNSLMGALCGFGLPIFYI 117

; CLONE: Consensus
US-09-726-899-10
; Sequence 10, Application US/09726899
; Publication No. US20010041356A1
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goli, Surya K.
; APPLICANT: Hillman, Jennifer L.
; TITLE OF INVENTION: NOVEL SUBUNITS OF NADH DEHYDROGENASE
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/726,899
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/785,065
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0187 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 129 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 114
; US-09-726-899-10

Query Match 75.8%; Score 514; DB 9; Length 129;
Best Local Similarity 73.6%; Pred. No. 3.3e-53;
Matches 95; Conservative 18; Mismatches 16; Indels 0; Gaps 0;

QY 1 MSFPKYPSSRLTLPETLDPAEYNISPETRAQAERLAIRAQLKREYLLQYNDPNRRGLI 60
Db 1 MSFPKYEASRLSLPTLDPAEYDISSETRKAQAERLAIRSLKREYQLQYDFSRGVI 60

QY 61 ENPALLRWAYARTINYPNFRPTPKNSLMGALCGFGLPIFYIYIINTERDRKELIQEGK 120
Db 61 EDPALVRWTYARSANIYPNFRPTPKNSLMGALCGFGLPIFYIYIINTERDRKELIQEGK 120

QY 121 LDRTHLSY 129
Db 121 LDRTHLSY 129

RESULT 4
US-10-106-698-4407
; Sequence 4407, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptides
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
```

```
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 4407
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-106-698-4407

Query Match
Best Local Similarity 18.3%; Score 124; DB 15; Length 33;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 106 KTERDRKKEKLIQEGKLDRTFHLIS 129
Db 10 KTERDRKKEKLIQEGKLDRTFHLIS 33

RESULT 5
US-09-925-299-1044
; Sequence 1044, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1044
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-299-1044

Query Match
Best Local Similarity 11.2%; Score 76; DB 9; Length 16;
Matches 15; Conservative 0; Mismatches 0; Indels 1; Gaps 0;

Qy 114 KLIQEGKLDRTFHLIS 129
Db 1 KLIQVGKLDRTFHLIS 16

RESULT 6
US-09-925-299-1044
; Sequence 1044, Application US/09925299
; Publication No. US20030040617A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1044
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Homo sapiens

US-09-925-299-1044

Query Match
Best Local Similarity 11.2%; Score 76; DB 11; Length 16;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 114 KLIQEGKLDRTFHLIS 129
Db 1 KLIQVGKLDRTFHLIS 16

RESULT 7
US-10-369-493-21273
; Sequence 21273, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 21273
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Xenorhabdus nematophilus
US-10-369-493-21273

Query Match
Best Local Similarity 10.5%; Score 71; DB 12; Length 433;
Matches 24; Conservative 29; Mismatches 37; Indels 28; Gaps 5;

Qy 10 SLRTLPTLPDAEYNISPETRAQAEALRAQLKREYLLQVNDPNRGLIENPALLRWA 69
Db 121 ALRTIPVLLDIH-----KVERLAPEA-----WIINFSNP--AGVV-TEAVSRYS 162
Qy 70 YARTI-----NVYPNFRPTPKNSLMGALCGFGLIFIYIYIKTERDRKELIQ 117
Db 163 KAKIIGLCNVFISPMHEHMANLQRPYQDVQLRFAGLNHMMVHVQVLNVRDETEQVLE 220

RESULT 8
US-10-369-493-2648
; Sequence 2648, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 2648
; LENGTH: 788
; TYPE: PRT
; ORGANISM: Synecocystis sp.
US-10-369-493-2648

Query Match
Best Local Similarity 10.4%; Score 70.5; DB 12; Length 788;
```

Best Local Similarity 30.5%; Pred. No. 25;
Matches 25; Conservative 9; Mismatches 33; Indels 15; Gaps 3;
QY 16 ETLDPAEYNISPETRAQALRAIAQLKREYLLQYNDPNRGLIENPALLRWAYARTI- 74
Db 597 EILDSLEKDDERDLRAQADLOQVLYELNREVRLOQDD-KEEGFFE-----AIKKTFT 638
QY 75 -----NVYPNFRPTPKNSLMG 90
Db 639 GDFDDDDYNNRRPAPRDDYRG 660

RESULT 9
US-10-032-585-7442
; Sequence 7442, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Bussey
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7442
; LENGTH: 936
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-032-585-7442

Query Match 10.4%; Score 70.5; DB 12; Length 936;
Best Local Similarity 29.2%; Pred. No. 31; Mismatches 14; Indels 25; Gaps 5;
Matches 26; Conservative 14; Mismatches 24; Indels 25; Gaps 5;
QY 6 YKPSLRTLPETLDPAEYNISPETRAQALRAIAQLKREYLL-----QYN-DPNRR 57
Db 118 YKPNARTLSKVLDPITVN-----AAERLFKNAIVDKNPFVSSAALISSYNLLPHAK 169
QY 58 GLIE---NPALLRWAYARTINYPNFRPT 83
Db 170 EVVKFTNETL-----ETIQSKSPPT 192

RESULT 10
US-09-814-777A-126
; Sequence 126, Application US/09814777A
; Patent No. US20020142415A1
; GENERAL INFORMATION:
; APPLICANT: KOOPMAN, Peter Anthony
; APPLICANT: MUSCAT, George Eugene Orlando
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND POLYNUCLEOTIDES AND METHODS OF USING THEM
; FILE REFERENCE: 21415-0003
; CURRENT APPLICATION NUMBER: US/09/814,777A
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: AU PQ6457
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 126
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Human
US-09-814-777A-126

Query Match 10.3%; Score 70; DB 10; Length 337;
Best Local Similarity 20.8%; Pred. No. 9.4; Mismatches 21; Indels 54; Gaps 5;
Matches 31; Conservative 21; Mismatches 43; Indels 54; Gaps 5;
QY 4 PKYKSSRLTLPETLDPAEYNISPETR--RAQALRLAIR-----AQLKREYLLQYN 52

Db 4 PASPPSPQSPSPSPGRLSPAGRGQAADSRIRPMPNAFWMVWAKDERKRLAQON 63
QY 53 D-----PNRRGLIENPALLRWAYARTINYPNFRPTPKNSLMGAL 92
Db 64 PDLHNAVLSKMLGKAWKELNAEKPPFVEEAERLVRVQHLD---HPNYKYRPR----- 113
QY 93 CGFGELIFIYIIKTERDRKEKLIQEGKL 121
Db 114 -----RKKQAKARRLEPGLL 129

RESULT 11
US-09-814-777A-15
; Sequence 15, Application US/09814777A
; Patent No. US20020142415A1
; GENERAL INFORMATION:
; APPLICANT: KOOPMAN, Peter Anthony
; APPLICANT: MUSCAT, George Eugene Orlando
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND POLYNUCLEOTIDES AND METHODS OF USING THE
; FILE REFERENCE: 21415-0003
; CURRENT APPLICATION NUMBER: US/09/814,777A
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: AU PQ6457
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 340
; TYPE: PRT
; ORGANISM: Human
US-09-814-777A-15

Query Match 10.3%; Score 70; DB 10; Length 340;
Best Local Similarity 20.8%; Pred. No. 9.5; Mismatches 21; Indels 54; Gaps 5;
Matches 31; Conservative 21; Mismatches 43; Indels 54; Gaps 5;
QY 4 PKYKSSRLTLPETLDPAEYNISPETR--RAQALRLAIR-----AQLKREYLLQYN 52
Db 4 PASPPSPQSPSPSPGRLSPAGRGQAADSRIRPMPNAFWMVWAKDERKRLAQON 63
QY 53 D-----PNRRGLIENPALLRWAYARTINYPNFRPTPKNSLMGAL 92
Db 64 PDLHNAVLSKMLGKAWKELNAEKPPFVEEAERLVRVQHLD---HPNYKYRPR----- 113
QY 93 CGFGELIFIYIIKTERDRKEKLIQEGKL 121
Db 114 -----RKKQAKARRLEPGLL 129

RESULT 12
US-09-814-777A-18
; Sequence 18, Application US/09814777A
; Patent No. US20020142415A1
; GENERAL INFORMATION:
; APPLICANT: KOOPMAN, Peter Anthony
; APPLICANT: MUSCAT, George Eugene Orlando
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND POLYNUCLEOTIDES AND METHODS OF USING THE
; FILE REFERENCE: 21415-0003
; CURRENT APPLICATION NUMBER: US/09/814,777A
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: AU PQ6457
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(482)
; NAME/KEY: misc.feature
; LOCATION: (679)..(1919)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 09:05:56 ; Search time 22 Seconds
(without alignments)
248.095 Million cell updates/sec

Title: US-09-726-899-3
Perfect score: 678
Sequence: 1 MSFPKYKPSRLTLPETLDP.....DRKEKLIQEKLDRTFHSY 129

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCFUS COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	678	100.0	129	2	US-08-785-065-3
2	678	100.0	129	4	US-09-151-412-3
3	514	75.8	129	2	US-08-785-065-10
4	514	75.8	129	4	US-09-151-412-10
5	79	11.7	345	4	US-09-252-991A-24156
6	70	10.3	360	4	US-08-252-991A-17420
7	69.5	10.3	294	2	US-08-874-347-26
8	69.5	10.3	294	3	US-09-093-522-26
9	69.5	10.3	354	4	US-09-252-991A-28779
10	69	10.2	640	4	US-09-592-595A-2
11	69	10.2	640	4	US-09-592-595A-4
12	69	10.2	656	3	US-08-738-000-4
13	69	10.2	656	3	US-09-258-928-4
14	69	10.2	656	4	US-09-347-878-24
15	69	10.2	656	4	US-09-660-872A-4
16	69	10.2	660	3	US-08-738-000-2
17	69	10.2	660	3	US-09-258-928-2
18	69	10.2	660	4	US-09-660-872A-2
19	68	10.0	196	4	US-09-227-357-393
20	68	10.0	228	4	US-09-227-357-401
21	68	10.0	371	4	US-09-199-637A-295
22	68	10.0	371	4	US-09-252-991A-21430
23	68	10.0	864	4	US-09-604-728-11
24	68	10.0	864	4	US-09-604-728-11
25	67.5	10.0	603	4	US-09-252-991A-31905
26	66.5	9.8	1036	4	US-09-252-991A-18349
27	66	9.7	855	4	US-09-252-991A-26493

28	65.5	9.7	969	4	US-09-252-991A-23580
29	65	9.6	1093	4	US-09-252-991A-21827
30	64.5	9.5	285	4	US-09-173-300-24
31	64.5	9.5	693	4	US-09-252-991A-20348
32	64	9.4	152	1	US-07-644-372-2
33	64	9.4	287	4	US-09-599-360B-95
34	64	9.4	305	2	US-08-946-528-1
35	63.5	9.4	686	4	US-09-944-259-33
36	63	9.3	553	4	US-09-252-991A-32970
37	63	9.3	1021	4	US-09-252-991A-19205
38	63	9.3	3898	3	US-08-750-717-2
39	62.5	9.2	286	4	US-09-252-991A-17236
40	62.5	9.2	366	2	US-08-605-106-11
41	62.5	9.2	375	4	US-09-328-352-7783
42	62	9.1	214	3	US-09-587-066-6
43	61.5	9.1	60	1	US-08-370-225-32
44	61.5	9.1	60	1	US-08-461-859-32
45	61.5	9.1	60	5	PCT-US93-10069-32

ALIGNMENTS

RESULT 1
US-08-785-065-3
; Sequence 3, Application US/08785065
; Patent No. 5814451
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goll, Surya K.
; APPLICANT: Hillman, Jennifer L.
; TITLE OF INVENTION: NOVEL SUBUNITS OF NADH DEHYDROGENASE
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304

COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/785,065
; FILING DATE: Herewith
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0187 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 129 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: Consensus
; CLONE: Consensus

US-08-785-065-3
Query Match 100.0%; Score 678; DB 2; Length 129;
Best Local Similarity 100.0%; Pred. No. 1.1e-74;

Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSFPKYKPSLRTLPETLDPAEYNISPTTRQAERLAIRAOQKREYLLQYNDPNRRGLI 60
 Db 1 MSFPKYKPSLRTLPETLDPAEYNISPTTRQAERLAIRAOQKREYLLQYNDPNRRGLI 60

QY 61 ENPALLWAYARTINVPNFRPTPKNSLMGALCGFGLPIFYIYIKTERDRKELIOEGK 120
 Db 61 ENPALLWAYARTINVPNFRPTPKNSLMGALCGFGLPIFYIYIKTERDRKELIOEGK 120

QY 121 LDRTFHLSY 129
 Db 121 LDRTFHLSY 129

RESULT 2

US-09-151-412-3
 ; Sequence 3, Application US/09151412
 ; Patent No. 6399345
 ; GENERAL INFORMATION:
 ; APPLICANT: Bandman, Olga
 ; APPLICANT: Goli, Surya K.
 ; APPLICANT: Hillman, Jennifer L.
 ; TITLE OF INVENTION: NOVEL SUBUNITS OF NADH DEHYDROGENASE
 ; NUMBER OF SEQUENCES: 12
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.
 ; STREET: 3174 Porter Drive
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94304

COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09151,412
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/785,065
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Billings, Lucy J.
 REGISTRATION NUMBER: 36,749
 REFERENCE/DOCKET NUMBER: PF-0187 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-855-0555
 TELEFAX: 415-845-4166
 TELEX:

INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 129 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 LIBRARY: Consensus
 CLONE: Consensus

US-09-151-412-3

Query Match 100.0%; Score 678; DB 4; Length 129;
 Best Local Similarity 100.0%; Pred. NO. 1.1e-74;
 Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSFPKYKPSLRTLPETLDPAEYNISPTTRQAERLAIRAOQKREYLLQYNDPNRRGLI 60
 Db 1 MSFPKYKPSLRTLPETLDPAEYNISPTTRQAERLAIRAOQKREYLLQYNDPNRRGLI 60

QY 61 ENPALLWAYARTINVPNFRPTPKNSLMGALCGFGLPIFYIYIKTERDRKELIOEGK 120

Db 61 ENPALLWAYARTINVPNFRPTPKNSLMGALCGFGLPIFYIYIKTERDRKELIOEGK 120

QY 121 LDRTFHLSY 129
 Db 121 LDRTFHLSY 129

RESULT 3

US-08-785-065-10
 ; Sequence 10, Application US/08785065
 ; Patent No. 5814451
 ; GENERAL INFORMATION:
 ; APPLICANT: Bandman, Olga
 ; APPLICANT: Goli, Surya K.
 ; APPLICANT: Hillman, Jennifer L.
 ; TITLE OF INVENTION: NOVEL SUBUNITS OF NADH DEHYDROGENASE
 ; NUMBER OF SEQUENCES: 12
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.
 ; STREET: 3174 Porter Drive
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94304

COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08785,065
 FILING DATE: Herewith
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Billings, Lucy J.
 REGISTRATION NUMBER: 36,749
 REFERENCE/DOCKET NUMBER: PF-0187 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-855-0555
 TELEFAX: 415-845-4166
 TELEX:

INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 129 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 LIBRARY: GenBank
 CLONE: 114

US-08-785-065-10

Query Match 75.8%; Score 514; DB 2; Length 129;
 Best Local Similarity 73.6%; Pred. No. 9e-55;
 Matches 95; Conservative 18; Mismatches 16; Indels 0; Gaps 0;

QY 1 MSFPKYKPSLRTLPETLDPAEYNISPTTRQAERLAIRAOQKREYLLQYNDPNRRGLI 60
 Db 1 MSFPKYKPSLRTLPETLDPAEYNISPTTRQAERLAIRAOQKREYLLQYNDPNRRGLI 60

QY 61 ENPALLWAYARTINVPNFRPTPKNSLMGALCGFGLPIFYIYIKTERDRKELIOEGK 120
 Db 61 EDPAVRYTYARSANIYVNFNRTKTSLLGALFGPLGVFWYVFKTDORKEKLIQEGK 120

QY 121 LDRTFHLSY 129

Db 121 LDRTFHLSY 129

RESULT 4

```
US-09-151-412-10
; Sequence 10, Application US/09151412
; Patent No. 6399345
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goli, Surya K.
; APPLICANT: Hillman, Jennifer L.
; TITLE OF INVENTION: NOVEL SUBUNITS OF NADH DEHYDROGENASE
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/151.412
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/785,065
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PP-0187 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 129 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 114
;
US-09-151-412-10
Query Match 75.8%; Score 514; DB 4; Length 129;
Best Local Similarity 73.6%; Pred. No. 9e-55;
Matches 95; Conservative 18; Mismatches 16; Indels 0; Gaps 0;
QY 1 MSPPKYFPSSLRKTLPTLDPAYNISPTETRAQAERLAIRAQLKREYLLQYNDPNRGLI 60
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 1 MSPPKYASRLSSLTPLDPAEYDISSETRAQAERLAIRSLKREYLOLYDPSRGGVI 60
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 61 ENPALLWAVARTNVVFNPRPTKNSLMGALCGFGLPIFYIYIKTERDRKEKLIQEGK 120
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 61 EDPALVWTVARSANIYFNPRNTKSLGALFGIGLVFWIYFKTDRDRKEKLIQEGK 120
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 121 LDRTFHLSY 129
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 121 LDRTFNISY 129
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |

RESULT 5
US-09-252-991A-24156
; Sequence 24156, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136

US-09-252-991A-24156
Query Match 11.7%; Score 79; DB 4; Length 346;
Best Local Similarity 37.5%; Pred. No. 0.22;
Matches 21; Conservative 5; Mismatches 22; Indels 8; Gaps 1;
QY 13 TLPETLDPAYNISPTETRAQAERLAIRAQLKREYLLQYNDPNRGLIENPALLEW 68
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 92 TLPEALDPA-----RRAGWRLACQCRVLSDLVLQFPDFPDGLPARVVACHW 139
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |

RESULT 6
US-09-252-991A-17420
; Sequence 17420, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17420
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
;
US-09-252-991A-17420
Query Match 10.3%; Score 70; DB 4; Length 360;
Best Local Similarity 26.0%; Pred. No. 2.9;
Matches 34; Conservative 17; Mismatches 52; Indels 28; Gaps 6;
QY 15 PETLDPAYNISPTETRAQAERLAIRAQLKREYLLQYNDPNRGLIENPALLRWAYAR-- 72
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 202 PPAIDPV-----SQRKAHAERVMARLQASIDLALQH-EPKSVVLGHGTGAYWA-ARVL 253
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 73 -----TINVYFNPRPTKNSLMGALCGFGLPIFYIYIYIKTERDRKEKLIQ 118
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 254 AEKEPABTHNLLVTAEVPRDPRPALEDWVPEKLATGD--FYRDARADREARLRMOA 311
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 119 GKLDRTFHSY 129
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 312 GK--RQKHPAY 320
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |

RESULT 7
US-08-874-347-26
; Sequence 26, Application US/08874347
; Patent No. 5863741
; GENERAL INFORMATION:
; APPLICANT: Limper, Andrew H.
; APPLICANT: Leof, Edward B.
; APPLICANT: Thomas, Charles F.
; APPLICANT: Gustafson, Michael P.
; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS
; FILE REFERENCE: 107196.136
```



```
Query Match      10.38; Score 69.5; DB 3; Length 294;
Best Local Similarity 36.2%; Pred.No.2.5;
Matches 17; Conservative 9; Mismatches 20; Indels
```

GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

TITLE OF INVENTION: NOVELLY ACID-BINDING AGENTS
 TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSIS
 FILE REFERENCE: 107196.136
 CURRENT APPLICATION NUMBER: US/09/252,991A
 CURRENT FILING DATE: 1999-02-18
 PRIOR APPLICATION NUMBER: US 60/074,788
 PRIOR FILING DATE: 1998-02-18
 PRIOR APPLICATION NUMBER: US 60/094,190
 PRIOR FILING DATE: 1998-07-27
 NUMBER OF SEQ ID NOS: 33142

```

; SEQ ID NO 28779
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28779

Query Match          10.3%; Score 69.5; DB 4; Length 354;
Best Local Similarity 30.1%; Pred. NO. 3.2;
Matches 28; Conservative 12; Mismatches 36; Indels 17; Gaps 5;

Qy      7  KPSSRLTLPETLDPAEYNI SPETRRQAERLATRAQLKREYLLQYNDPNRRGLIENPALL 66
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      142 RPADLPRLP--VDMSRHGGVPGNRRDLRRV-VRAGLR-----GDPRR-----DPAAH 187

Qy      67  RWAYARTINYPNFRPTPKNSLMGALCGFGPLI 99
Db      188 RPKOGAD----DPRHPGPGAGNLCAGAGGGSPAV 217

```

```

RESULT 10
US-09-592-595A-2
; Sequence 2, Application US/0952595A
; Patent No. 6566065
; GENERAL INFORMATION:
; APPLICANT: ROZEN, Rima
; TITLE OF INVENTION: CDNA FOR HUMAN METHYLE
; TITLE OF INVENTION: REDUCTASE AND USES TH
; FILE REFERENCE: 04844/005003
; CURRENT APPLICATION NUMBER: US/09/592,595A
; CURRENT FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: US 09/258,928

```


QY 35 ERLAIRAQLKREYLLQYNDPNRRGLIENPALLRWAYARTINVYPNFRPTPKNSLMGALCG 94
 Db 458 EPLAARTSLKELLRV---NRQIL-----TINSQPNINGKPPSSD---PIVG 499
 QY 95 FGP---LIP---IYYIKTERDRKEKLIQ---BGKLDRTFHL 127
 Db 500 WGPSSGYVFKAYLEFPTSRETAELQLVKKYELRVNYHL 540

RESULT 14
 US-09-347-878-24
 ; Sequence 24, Application US/09347878C
 ; Patent No. 6376210
 ; GENERAL INFORMATION:
 ; APPLICANT: Yuan, Chong
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR ASSAYING ANALYTES
 ; FILE REFERENCE: 25885-1651
 ; CURRENT APPLICATION NUMBER: US/09/347,878C
 ; CURRENT FILING DATE: 1999-07-06
 ; NUMBER OF SEQ ID NOS: 75
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 24
 ; LENGTH: 656
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: Human methylenetetrahydrofolate reductase (MTHFR)
 ; OTHER INFORMATION: protein sequence
 ; PUBLICATION INFORMATION:
 ; DATABASE ACCESSION NUMBER: AF105977/GenBank 1-11
 US-09-347-878-24

Query Match 10.2%; Score 69; DB 4; Length 656;
 Best Local Similarity 28.7%; Pred. No. 8.5; 32; Indels 26; Gaps 6;
 Matches 29; Conservative 14; Mismatches 32; Indels 26; Gaps 6;
 QY 35 ERLAIRAQLKREYLLQYNDPNRRGLIENPALLRWAYARTINVYPNFRPTPKNSLMGALCG 94
 Db 458 EPLAARTSLKELLRV---NRQIL-----TINSQPNINGKPPSSD---PIVG 499
 QY 95 FGP---LIP---IYYIKTERDRKEKLIQ---BGKLDRTFHL 127
 Db 500 WGPSSGYVFKAYLEFPTSRETAELQLVKKYELRVNYHL 540

RESULT 15
 US-09-660-872A-4
 ; Sequence 4, Application US/09660872A
 ; Patent No. 6528259
 ; GENERAL INFORMATION:
 ; APPLICANT: ROZEN, Rima
 ; APPLICANT: GOYETTE, Philippe
 ; TITLE OF INVENTION: CDNA FOR HUMAN METHYLENETETRAHYDROFOLATE
 ; TITLE OF INVENTION: REDUCTASE
 ; FILE REFERENCE: 04844/005004
 ; CURRENT APPLICATION NUMBER: US/09/660,872A
 ; CURRENT FILING DATE: 2000-09-13
 ; PRIOR APPLICATION NUMBER: PCT/IB00/00442
 ; PRIOR FILING DATE: 2000-02-28
 ; PRIOR APPLICATION NUMBER: US 09/258,928
 ; PRIOR FILING DATE: 1999-03-01
 ; PRIOR APPLICATION NUMBER: US 08/738,000
 ; PRIOR FILING DATE: 1997-02-12
 ; PRIOR APPLICATION NUMBER: PCT/CA95/00314
 ; PRIOR FILING DATE: 1995-05-25
 ; PRIOR APPLICATION NUMBER: GB 9410620.0
 ; PRIOR FILING DATE: 1994-05-26
 ; NUMBER OF SEQ ID NOS: 18
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 4
 ; LENGTH: 656
 ; TYPE: PRT

; ORGANISM: Homo sapiens
 US-09-660-872A-4
 Query Match 10.2%; Score 69; DB 4; Length 656;
 Best Local Similarity 28.7%; Pred. No. 8.5;
 Matches 29; Conservative 14; Mismatches 32; Indels 26; Gaps 6;
 QY 35 ERLAIRAQLKREYLLQYNDPNRRGLIENPALLRWAYARTINVYPNFRPTPKNSLMGALCG 94
 Db 458 EPLAARTSLKELLRV---NRQIL-----TINSQPNINGKPPSSD---PIVG 499
 QY 95 FGP---LIP---IYYIKTERDRKEKLIQ---BGKLDRTFHL 127
 Db 500 WGPSSGYVFKAYLEFPTSRETAELQLVKKYELRVNYHL 540
 Search completed: January 29, 2004, 09:09:35
 Job time : 23 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 09:14:07 ; Search time 20 Seconds
(without alignments)
620.288 Million cell updates/sec

Title: US-09-726-899-3

Perfect score: 129

Sequence: 1 MSPFKYKPSLRLTPETLDP.....DRKEKLIQEGKGLDRTFHLIS 129

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283308 seqs, 96168682 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 75 summaries

Database :

PIR.76:*

1: Pirl1:*

2: Pirl2:*

3: Pirl3:*

4: Pirl4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	129	100.0	129	2	J80383
2	17	13.2	129	2	S28237
3	8	6.2	133	2	JC2003
4	8	6.2	160	2	B71176
5	8	6.2	179	2	A33164
6	8	6.2	244	2	A31318
7	8	6.2	244	2	A31690
8	8	6.2	451	2	H82044
9	8	6.2	466	2	E95262
10	8	6.2	468	2	E83046
11	8	6.2	658	2	T40107
12	8	6.2	1440	1	SVHUOT
13	7	5.4	127	2	AH0316
14	7	5.4	227	2	T22144
15	7	5.4	239	2	A00088
16	7	5.4	252	2	D95266
17	7	5.4	278	2	C41841
18	7	5.4	278	2	A01872
19	7	5.4	282	2	AC2047
20	7	5.4	310	2	T36068
21	7	5.4	321	2	G82688
22	7	5.4	347	2	T31205
23	7	5.4	366	2	C87624
24	7	5.4	398	2	F90206
25	7	5.4	400	2	T27952
26	7	5.4	438	2	G96035
27	7	5.4	587	2	I40849
28	7	5.4	608	2	A64992
29	7	5.4	628	2	S77374

30	7	5.4	681	2	G87276	hypothetical prote
31	7	5.4	748	2	I48744	semaphorin A - mou
32	7	5.4	967	2	H71087	leucine-trRNA ligas
33	7	5.4	967	2	H75133	leucyl-tRNA synthe
34	7	5.4	1111	2	T01078	hypothetical prote
35	6	4.7	29	1	GCFLF	glucagon - Europea
36	6	4.7	29	2	A61135	glucagon - bigeye
37	6	4.7	29	2	C60840	glucagon I - Europ
38	6	4.7	57	2	C48642	hypothetical prote
39	6	4.7	58	2	P00129	34.5K linker prote
40	6	4.7	60	2	A02255	hypothetical prote
41	6	4.7	63	1	GCIDC	glucagon precursor
42	6	4.7	78	2	AC2214	hypothetical prote
43	6	4.7	80	2	B23617	conglutin delta-2
44	6	4.7	87	1	GCFFS	glucagon precursor
45	6	4.7	89	2	D90802	probable transcrip
46	6	4.7	93	2	D59960	hypothetical prote
47	6	4.7	97	2	S76406	hypothetical prote
48	6	4.7	99	2	A69514	hypothetical prote
49	6	4.7	108	2	T23220	hypothetical prote
50	6	4.7	117	2	F95862	conserved hypotet
51	6	4.7	118	2	B86787	hypothetical prote
52	6	4.7	122	1	GCAP2	glucagon 2 precurs
53	6	4.7	123	2	S78256	ribosomal protein
54	6	4.7	123	2	S18465	matrix protein M2
55	6	4.7	125	2	F84604	hypothetical prote
56	6	4.7	133	2	T36401	hypothetical prote
57	6	4.7	142	2	AG2088	hypothetical prote
58	6	4.7	145	2	E82573	hypothetical prote
59	6	4.7	153	2	A33090	conglutin delta pr
60	6	4.7	155	2	T48503	hypothetical prote
61	6	4.7	158	2	E72702	hypothetical prote
62	6	4.7	159	2	AC1103	7,8-dihydro-6-hydr
63	6	4.7	159	2	B24691	34.5K linker polyp
64	6	4.7	164	2	D42148	GTP-binding protei
65	6	4.7	171	2	E84912	probable polyribon
66	6	4.7	173	2	H87629	conserved hypotet
67	6	4.7	177	2	G75285	hypothetical prote
68	6	4.7	180	2	S73221	ribosomal protein
69	6	4.7	181	2	S35485	resolvase - Pseudo
70	6	4.7	182	2	J01753	hypothetical 22.3K
71	6	4.7	183	2	G75019	hami protein relat
72	6	4.7	183	2	T29987	hypothetical prote
73	6	4.7	186	2	G71725	ribosome recycling
74	6	4.7	186	2	E97724	cell division prot
75	6	4.7	191	2	B69506	

ALIGNMENTS

RESULT 1

JE0383

NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain NDUFB4 - human

N:Alternate names: NADH:ubiquinone oxidoreductase subunit NDUFB4

C:Species: Homo sapiens (man)

C>Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 03-Jun-2002

C:Accession: JE0383

R:Loeffen, J.L.C.M.; Triepels, R.H.; van den Heuvel, L.P.; Schuelke, M.; Buskens, C.A.)

Biochem. Biophys. Res. Commun. 253, 415-422, 1998

A:Title: cDNA of eight nuclear encoded subunits of NADH:ubiquinone oxidoreductase: Hum

A:Reference number: JE0379; MUID:99097250; PMID:9878551

A:Accession: JE0383

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-129 <LOE>

A:Cross-references: GB:AF044957; NID:G4164445; PIDN:AA05421.1; PID:G4164446

C:Keywords: NAD; oxidoreductase

Query Match 100.0%; Score 129; DB 2; Length 129;

Best Local Similarity 100.0%; Pred. No. 2.5e-126;

Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSFPKYKSSSLTLPTLDPAYNISPETTRQAERLAIRAQLKREYLLQYNDPNRRGLI 60
DB 1 MSFPKYKSSSLTLPTLDPAYNISPETTRQAERLAIRAQLKREYLLQYNDPNRRGLI 60
QY 61 ENPALLWAYARTINYPNFRTPKNSLGMALCGGFLPIFYIYIITKTERDRKEKLIQEGK 120
DB 61 ENPALLWAYARTINYPNFRTPKNSLGMALCGGFLPIFYIYIITKTERDRKEKLIQEGK 120
QY 121 LDRTHLSY 129
DB 121 LDRTHLSY 129

RESULT 2

S28237
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain B15 - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 03-Jun-2002
C:Accession: S28237
R:Walker, J.E.; Arizmendi, J.M.; Dupuis, A.; Fearnley, I.M.; Finel, M.; Medd, S.M.; Pille
J. Mol. Biol. 226, 1051-1072, 1992
A:Title: Sequences of 20 subunits of NADH: ubiquinone oxidoreductase from bovine heart
A:Reference number: S28237; MUID:92389317; PMID:1518044
A:Accession: S28237
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-129 <WAL>
A:Cross-references: EMBL:X64898; NID:g113; PIDN:CAA46107.1; PID:g114
C:Keywords: electron transfer; mitochondrion; NAD; oxidoreductase

Query Match 13.2%; Score 17; DB 2; Length 129;
Best Local Similarity 100.0%; Pred. No. 5.4e-10;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 109 RDRKEKLIQEGKLDRTF 125
DB 109 RDRKEKLIQEGKLDRTF 125

RESULT 3

JC2003
NADH ubiquinone oxidoreductase B15 chain like protein - chicken
N:Alternate names: murine Hox-3.1 homeobox
C:Species: Gallus gallus (chicken)
C:Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 07-Feb-1997
C:Accession: JC2003
R:Goldberg, G.S.; Kaczmarczyk, W.
Gene 133, 233-235, 1993
A:Title: A chicken genomic DNA fragment that hybridizes to the murine Hox-3.1 homeobox
A:Reference number: JC2003; MUID:94040816; PMID:7901127
A:Accession: JC2003
A:Molecule type: DNA
A:Residues: 1-133 <GOL>
C:Genetics:
A:Gene: gghpw
A:Introns: 67/3
C:Keywords: homeobox; phosphoprotein; sulfoxoprotein; transmembrane protein
F:95-112/Domain: transmembrane #status predicted <TM>
F:30/Binding site: sulfate (Tyr) (covalent) #status predicted
F:114/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted

Query Match 6.2%; Score 8; DB 2; Length 133;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 AERLAIRA 41
DB 41 AERLAIRA 48

RESULT 4

B71176
hypothetical protein PH1689 - Pyrococcus horikoshii

C:Species: Pyrococcus horikoshii
C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 21-Jul-2000
C:Accession: B71176
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sek
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguc
DNA Res. 5, 55-76, 1998
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic
A:Reference number: A71000; MUID:98344137; PMID:9679194
A:Accession: B71176
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-160 <KAW>

A:Cross-references: GB:AP000007; NID:g3236134; PIDN:BAA30801.1; PID:g3258118

A:Experimental source: strain OT3
A:Note: this accession replaces an interim accession for a sequence replaced by GenBan
C:Genetics:
A:Gene: PH1689

C:Superfamily: Pyrococcus horikoshii hypothetical protein PH1689

Query Match 6.2%; Score 8; DB 2; Length 160;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 KPSSSLRTL 14
DB 75 KPSSSLRTL 82

RESULT 5

A3164
hypothetical protein walter - chicken
N:Alternate names: homeotic protein Hox-3.1 homolog; hypothetical 13.1K protein
C:Species: Gallus gallus (chicken)
C:Date: 28-Mar-1991 #sequence_revision 28-Mar-1991 #text_change 21-Jul-2000
C:Accession: A3164; JCI395
R:Goldberg, G.S.
submitted to the Protein Sequence Database, March 1991
A:Reference number: A3164

A:Status: preliminary; not compared with conceptual translation

A:Accession: A3164

A:Molecule type: DNA

A:Residues: 1-179 <GOL>

A:Experimental source: strain Leghorn

R:Goldberg, G.S.; Kaczmarczyk, W.

Gene 121, 397-398, 1992

A:Title: Sequence of a novel chicken genomic DNA fragment that hybridizes to the murin

A:Reference number: JCI395; MUID:9307061; PMID:1359990

A:Accession: JCI395

A:Molecule type: DNA

A:Residues: 1-124 <GO2>

A:Cross-references: GB:M84354; NID:g211948; PIDN:AAA70193.1; PID:g211949

Query Match 6.2%; Score 8; DB 2; Length 179;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 AERLAIRA 41
DB 55 AERLAIRA 62

RESULT 6

AE1318
hypothetical protein lmc1949 [imported] - Listeria monocytogenes (strain EGD-e)
C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C:Accession: AE1318
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeck
; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussauguet, O.; Entian, K.D.; Fsihi,
D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.;
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlan

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 09:12:47 ; Search time 17 Seconds

(without alignments)
356,850 Million cell updates/sec

Title: US-09-726-899-3

Perfect score: 129

Sequence: 1 MSFPKYPSSRLTLPETLDP.....DRKEKLIQEGKLDRTFFHLSY 129

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 127863 seqs, 47026705 residues

Word size : 0

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 75 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	128	99.2	128	1	NBSM_HUMAN
2	17	13.2	128	1	NBSM_BOVIN
3	8	6.2	133	1	NBSM_CHICK
4	8	6.2	466	1	SELA_RHIME
5	8	6.2	468	1	SELA_PSEAE
6	8	6.2	1440	1	SYEP_HUMAN
7	7	5.4	127	1	CRCB_YERPE
8	7	5.4	277	1	PYR2_ANASP
9	7	5.4	398	1	HISX_SULSO
10	7	5.4	532	1	NADB_RHIME
11	7	5.4	536	1	CAR9_HUMAN
12	7	5.4	577	1	DCPY_ASPPA
13	7	5.4	587	1	DHSA_COXBU
14	7	5.4	608	1	ATOS_ECOLI
15	7	5.4	628	1	FTH3_SYNY3
16	7	5.4	748	1	SM3B_MOUSE
17	7	5.4	967	1	SVL_PYRAB
18	7	5.4	967	1	SVL_PYRHO
19	7	5.4	29	1	GLUC_PLAPE
20	6	4.7	57	1	YAC2_LEGPN
21	6	4.7	71	1	GLUC_ICTPU
22	6	4.7	80	1	GLUC_LUPAN
23	6	4.7	93	1	YOHV_BACSU
24	6	4.7	96	1	GLUC_MYOSC
25	6	4.7	99	1	YL13_ARCFU
26	6	4.7	117	1	VATG_DROME
27	6	4.7	121	1	GLUC_CARAU
28	6	4.7	122	1	GLU2_LOPAM
29	6	4.7	123	1	RL13_ODOSI
30	6	4.7	123	1	VMAT_VHSVO
31	6	4.7	135	1	NUSB_CLOPE
32	6	4.7	144	1	AROQ_RHILO
33	6	4.7	150	1	SP18_DROME
					Q95168 homo sapien
					P48305 bos taurus
					P48306 gallus gall
					P58226 rhizobium m
					Q9hv01 pseudomonas
					P07814 homo sapien
					Q82dh2 yersinia pe
					P31329 anabaena sp
					O33775 sulfolobus
					Q92i32 rhizobium m
					Q9h257 homo sapien
					P51844 aspergillus
					P51054 coxiella bu
					Q06067 escherichia
					P73437 synechocyst
					Q62177 mus musculu
					Q9v0b9 pyrococcus
					O58698 pyrococcus
					P23062 platicthys
					P37034 legionella
					P04093 ictalurus p
					P09331 lupinus ang
					P49779 bacillus su
					P09686 myxocephal
					O28167 archaeoglob
					Q9xxh6 drosophila
					P79695 carassius a
					P04092 lophius ame
					P49501 odontella s
					P27663 viral hemor
					Q8xjd6 clostridium
					Q98nc1 rhizobium l
					Q9vex9 drosophila

ALIGNMENTS

RESULT 1
NBSM_HUMAN
ID NBSM_HUMAN STANDARD; PRT; 128 AA.
AC O95168;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE NADH-ubiquinone oxidoreductase B15 subunit (EC 1.6.5.3) (EC 1.6.99.3)
DE (Complex I-B15) (CI-B15).
GN NDUF84.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99077250; PubMed=9878551;
RA Loeffen J.I.C.M., Triepels R.H., van den Heuvel L., Schuelke M.,
RA Buskens C.A.F., Smeets R.J.P., Trijbels J.M.F., Smeitink J.A.M.;
RT "cDNA of eight nuclear encoded subunits of NADH:ubiquinone
RT oxidoreductase: human complex I cDNA characterization completed.";
RN Biochem. Biophys. Res. Commun. 253:415-422(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=Cervix;
RX MEDLINE=22388257; PubMed=12477932;

P55547 rhizobium s
P35286 rattus norv
Q9ma36 arabidopsis
P51300 porphyra pu
P30739 pseudomonas
P30772 rickettsia
Q92e08 rickettsia
P58833 pyrococcus
Q12690 saccharomyc
P40212 saccharomyc
P39100 escherichia
Q00744 mus musculu
Q21985 caenorhabdi
P29960 rhodobacter
P15940 bradyrhizob
O07650 campylobact
O42787 metarhizium
O22288 arabidopsis
P58644 yersinia pe
O53899 mycobacteri
Q9hy05 pseudomonas
Q59188 borrelia bu
Q9hv51 pseudomonas
P95576 pseudomonas
P46106 streptomyce
P23138 rhodospiril
P35039 anopheles g
P42919 bacillus su
P11399 mastigoclad
P08501 rhodobacter
P73242 synechocyst
Q9bq99 homo sapien
O58137 methanococc
P81376 plasmodium
Q9zm13 helicobacte
O25657 helicobacte
Q61450 mus musculu
O07402 mycobacteri
Q27436 methanobact
Q8ew81 mycoplasma
Q8dvm7 streptococc

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altschul S.F., Buetow K.H., Schaefer C.P., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalak U., Smailus D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: TRANSFER OF ELECTRONS FROM NADH TO THE RESPIRATORY
 CC CHAIN. THE IMMEDIATE ELECTRON ACCEPTOR FOR THE ENZYME IS BELIEVED
 CC TO BE UBIQUINONE.
 CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
 CC -1- CATALYTIC ACTIVITY: NADH + acceptor = NAD(+) + reduced acceptor.
 CC -1- SUBUNIT: COMPLEX I IS COMPOSED OF ABOUT 40 DIFFERENT SUBUNITS.
 CC -1- SUBCELLULAR LOCATION: Mitochondrial inner membrane; matrix side.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF044957; AAD05421.1; -;
 CC EMBL; BC000855; AAH00855.1; -;
 CC PIR; JE0383; JE0383.
 CC Genew; HGNC:7699; NDUFB4.
 CC MIM; 603840; -;
 CC GO; GO:0005748; C:NADH dehydrogenase complex (ubiquinone) (se. .; TAS.
 CC GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; TAS.
 CC GO; GO:0006120; P:oxidative phosphorylation, NADH to ubiquinone; TAS.
 CC Oxidoreductase; NAD; Ubiquinone; Mitochondrion; Acetylation.
 CC INIT MET 0 0 BY SIMILARITY.
 CC MOD RES 1 1 ACETYLATION (BY SIMILARITY).
 CC SEQUENCE 128 AA; 15077 MW; DCCFG6BF646F99D8 CRC64;
 CC
 CC Query Match 99.2%; Score 128; DB 1; Length 128;
 CC Best Local Similarity 100.0%; Pred. No. 1.7e-126; Indels 0; Gaps 0;
 CC Matches 128; Conservative 0; Mismatches 0;
 QY 2 SPFKYFSSLTLPETLDPAYNISPTERRAQLRAQLKREYLLQYNDPNERGLIE 61
 DB 1 SPFKYFSSLTLPETLDPAYNISPTERRAQLRAQLKREYLLQYNDPNERGLIE 60
 QY 62 NPALLWAYARTINYNFPTPKNSLGMALCGFGLPIFYIYIKTERDRKELIQGKL 121
 DB 61 NPALLWAYARTINYNFPTPKNSLGMALCGFGLPIFYIYIKTERDRKELIQGKL 120
 QY 122 DRTFHLST 129
 DB 121 DRTFHLST 128
 RESULT 2
 NBSM_BOVIN STANDARD; PRT; 128 AA.
 ID NBSM_BOVIN
 AC P48305;
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 RT "A chicken genomic DNA fragment that hybridizes to the murine Hox-3.1
 homeobox is likely to encode the NADH ubiquinone oxidoreductase
 subunit B15.";
 RT subunit B15.";

DE NADH-ubiquinone oxidoreductase B15 subunit (EC 1.6.5.3) (EC 1.6.99.3)
 DE (Complex I-B15) (CI-B15).
 GN NDUFB4.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Eubovae; Bos.
 RN NCBI_TaxID=9913;
 RX [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Heart;
 RC MEDLINE=92389317; PubMed=1518044;
 RA Walker J.E., Arizumi J.M., Dupuis A., Pearnley I.M., Finel M.,
 RA Medd S.M., Pilkington S.J., Runswick M.J., Skehel J.M.;
 RT "Sequences of 20 subunits of NADH:ubiquinone oxidoreductase from
 RT bovine heart mitochondria. Application of a novel strategy for
 RT sequencing proteins using the polymerase chain reaction.";
 RL J. Mol. Biol. 226:1051-1072(1992).
 CC -1- FUNCTION: TRANSFER OF ELECTRONS FROM NADH TO THE RESPIRATORY
 CC CHAIN. THE IMMEDIATE ELECTRON ACCEPTOR FOR THE ENZYME IS BELIEVED
 CC TO BE UBIQUINONE.
 CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
 CC -1- CATALYTIC ACTIVITY: NADH + acceptor = NAD(+) + reduced acceptor.
 CC -1- SUBUNIT: COMPLEX I IS COMPOSED OF ABOUT 40 DIFFERENT SUBUNITS.
 CC -1- SUBCELLULAR LOCATION: Mitochondrial inner membrane; matrix side.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X64898; CAA46107.1; -;
 CC PIR; S28237; S28237.
 CC Oxidoreductase; NAD; Ubiquinone; Mitochondrion; Acetylation.
 CC INIT MET 0 0
 CC MOD RES 1 1 ACETYLATION.
 CC SEQUENCE 128 AA; 15053 MW; CC1352E9E80DF7D5 CRC64;
 CC
 CC Query Match 13.2%; Score 17; DB 1; Length 128;
 CC Best Local Similarity 100.0%; Pred. No. 2e-10;
 CC Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 109 RDRKEXLIQEGKLDRTF 125
 DB 108 RDRKEXLIQEGKLDRTF 124
 RESULT 3
 NBSM_CHICK STANDARD; PRT; 133 AA.
 ID NBSM_CHICK
 AC P48306;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Putative NADH-ubiquinone oxidoreductase B15 subunit (EC 1.6.5.3)
 DE (EC 1.6.99.3) (Complex I-B15) (Hypothetical protein Walter)
 DE (GGHPW).
 GN NDUFB4.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 RN NCBI_TaxID=9031;
 RX [1]
 RP SEQUENCE FROM N.A.
 RP MEDLINE=94040816; PubMed=7901127;
 RA Goldberg G.S., Kaczmarczyk W.;
 RT "A chicken genomic DNA fragment that hybridizes to the murine Hox-3.1
 homeobox is likely to encode the NADH ubiquinone oxidoreductase
 subunit B15.";
 RT subunit B15.";

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 09:13:22 ; Search time 35 Seconds

(without alignments)
951.108 Million cell updates/sec

Title: US-09-726-899-3

Perfect score: 129

Sequence: 1 MSPFKYKPSLRTLPETLDP.....DRKEKLIQEGKLDRTFHLVS 129

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 830525 seqs, 258052604 residues

Word size : 0

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 75 summaries

Database :

SPTREMBL_23.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_xvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	108	83.7	119	4 Q8N4D3	Q8b4d3 homo sapien
2	12	9.3	119	11 Q9DHE2	Q8b4d3 mus musculu
3	12	9.3	129	11 Q9CQC7	Q8cqc7 mus musculu
4	8	6.2	160	17 Q59313	Q59313 pyrococcus
5	8	6.2	244	16 Q92A59	Q92a59 listeria in
6	8	6.2	244	16 Q8Y5V6	Q8y5v6 listeria mo
7	8	6.2	345	17 Q8TSM5	Q8tsm5 methanosarc
8	8	6.2	435	16 Q8DCY9	Q8dcy9 vibrio vuln
9	8	6.2	451	16 Q9KNN2	Q9knn2 schizosacch
10	8	6.2	658	3 Q74798	Q74798 schizosacch
11	8	6.2	1022	11 Q8CGC7	Q8cgc7 mus musculu
12	7	5.4	88	16 Q8Y2S6	Q8y2s6 ralstonia s
13	7	5.4	102	11 Q8C3A8	Q8c3a8 mus musculu
14	7	5.4	120	5 Q8T539	Q8t539 plasmodium
15	7	5.4	120	5 Q77180	Q77180 plasmodium
16	7	5.4	174	11 Q8BP01	Q8bp01 mus musculu

RESULT 1

ID	Q8N4D3	PRELIMINARY;	PRT;	119 AA.
AC	Q8N4D3	01-OCT-2002 (Tremblrel. 22, Created)		
DT	01-OCT-2002 (Tremblrel. 22, Last sequence update)			
DE	01-OCT-2002 (Tremblrel. 22, Last annotation update)			
DE	Hypothetical protein (Fragment)			
OS	Homo sapiens (Human)			

ALIGNMENTS

17	5.4	180	12	Q8V5P7	Q8v5p7 helicoverpa
18	5.4	181	12	Q99G88	Q99g88 helicoverp
19	5.4	181	12	Q91BS9	Q91bs9 helicoverpa
20	5.4	194	16	Q9RJL5	Q9rjl5 streptomyce
21	5.4	227	5	Q93716	Q93716 caenorhabdi
22	5.4	239	16	Q8Z106	Q8z106 versinia pe
23	5.4	247	16	Q8CKH8	Q8ckh8 versinia pe
24	5.4	252	16	Q931A8	Q931a8 rhizobium m
25	5.4	260	5	Q8SW63	Q8sw63 encephalito
26	5.4	276	9	Q8SD96	Q8sd96 pseudomonas
27	5.4	282	16	Q8YVP5	Q8yvp5 anabaena sp
28	5.4	301	17	Q8TPD8	Q8tpd8 methanosarc
29	5.4	310	16	Q9ZBX0	Q9zbx0 streptomyce
30	5.4	321	16	Q9PDK0	Q9pdk0 xyella fas
31	5.4	341	10	Q9LWK8	Q9lwk8 oryza sativ
32	5.4	347	2	Q85914	Q85914 sphingomona
33	5.4	362	10	Q9FPJ8	Q9fpj8 arabidopsis
34	5.4	366	16	Q9A414	Q9a414 caulobacter
35	5.4	400	5	Q23568	Q23568 caenorhabdi
36	5.4	438	16	Q92TG8	Q92tg8 rhizobium m
37	5.4	506	5	Q8IJJ7	Q8ijj7 plasmodium
38	5.4	514	10	Q8LPD0	Q8lpd0 oryza sativ
39	5.4	538	16	Q8REJ1	Q8rej1 fusobacteri
40	5.4	556	5	Q8I610	Q8i610 plasmodium
41	5.4	576	5	Q8WRE3	Q8wre3 anopheles g
42	5.4	612	16	Q8FFP8	Q8ffp8 escherichia
43	5.4	619	16	Q8DLG5	Q8dlg5 synecococc
44	5.4	627	4	Q60826	Q60826 homo sapien
45	5.4	627	11	Q9JIG7	Q9jig7 mus musculu
46	5.4	627	11	Q8BYH4	Q8byh4 mus musculu
47	5.4	636	3	Q9Y8A4	Q9y8a4 aspergillus
48	5.4	678	10	Q9LHN5	Q9lhn5 arabidopsis
49	5.4	681	16	Q9ABK3	Q9abk3 caulobacter
50	5.4	699	16	Q98JPS	Q98jp5 rhizobium 1
51	5.4	701	5	Q95T82	Q95t82 drosophila
52	5.4	793	5	Q9VDV4	Q9vuv4 drosophila
53	5.4	857	10	Q9LNM4	Q9lnm4 arabidopsis
54	5.4	875	5	Q9V6A2	Q9v6a2 drosophila
55	5.4	896	5	Q8MRB3	Q8mrb3 drosophila
56	5.4	1111	10	Q9ZQX8	Q9zqx8 arabidopsis
57	5.4	1179	5	Q8IKC3	Q8ikc3 plasmodium
58	5.4	2208	16	Q8XRC8	Q8xrc8 ralstonia s
59	5.4	2291	5	Q9W1A9	Q9wia9 drosophila
60	5.4	3170	2	Q9ALM4	Q9alm4 saccharopol
61	5.4	3347	5	Q8IEA1	Q8iea1 plasmodium
62	5.4	3513	2	Q8GBX5	Q8gbx5 polyangium
63	5.4	3524	2	Q93H86	Q93h86 streptomyce
64	5.4	4150	2	Q9KIV4	Q9kiv4 streptomyce
65	5.4	4557	16	Q8CJN6	Q8cjn6 streptomyce
66	5.4	5017	2	Q8GBX6	Q8gbx6 polyangium
67	5.4	6048	2	Q93H87	Q93h87 streptomyce
68	5.4	6145	2	Q93H84	Q93h84 streptomyce
69	4.7	11	6	Q9TQS9	Q9tqs9 equus cabal
70	4.7	20	2	Q9R636	Q9r636 desulfovibr
71	4.7	36	16	Q8EY79	Q8ey79 leptospira
72	4.7	53	6	Q9GL22	Q9gl22 canis fami
73	4.7	54	4	Q9H256	Q9h256 homo sapien
74	4.7	60	16	Q8YR55	Q8yr55 anabaena sp
75	4.7	69	1	Q9C4Y2	Q9c4y2 sulfolobus


```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1] NCBI_TaxID=9606;
RP SEQUENCE FROM N.A.
RC STRAUSBERG R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC034579; AAH34579.1; -
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 119 AA; 13921 MW; 38EAPFA46S5AE09BA CRC64;

Query Match 83.7%; Score 108; DB 4; Length 119;
Best Local Similarity 100.0%; Pred. No. 7.3e-105;
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SPKYPSSLRITLPTLDPAYNISPTETRAQERLAIRALQKREYLLQYNDPDRGLIE 61
DB 1 SPKYPSSLRITLPTLDPAYNISPTETRAQERLAIRALQKREYLLQYNDPDRGLIE 60

QY 62 NPALLRWARTINVTNPTKNSLMGALCGFGLPIFYIYIKTER 109
DB 61 NPALLRWARTINVTNPTKNSLMGALCGFGLPIFYIYIKTER 108

RESULT 2
Q9DBH2
ID Q9DBH2 PRELIMINARY; PRT; 119 AA.
AC Q9DBH2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE 1300010H20Rik protein.
GN 1300010H20Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Liver;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaiji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR ENBL; AK018707; BAB31358.1; -
DR MGD; MGI:1915444; 0610006N12Rik.
SQ SEQUENCE 129 AA; 15081 MW; 9E76B5087F095062 CRC64;

Query Match 9.3%; Score 12; DB 11; Length 129;
Best Local Similarity 100.0%; Pred. No. 0.00022;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 LKREYLLQYNDP 54
DB 43 LKREYLLQYNDP 54

RESULT 4
O59313 PRELIMINARY; PRT; 160 AA.
ID O59313;
AC O59313;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DE 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein PH1689.
GN PH1689.
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=53953;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OT3;
RX MEDLINE=98344137; PubMed=9679194;
RA Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 09:09:42 ; Search time 41 Seconds

(without alignments)
499.408 Million cell updates/sec

Title: US-09-726-899-3

Perfect score: 129

Sequence: 1 MSPFKYKPSRLTLPETLDP.....DRKEKLQEGKLDRTFHLST 129

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 1107863 seqs, 158726573 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 75 summaries

Database :

A Geneseq 19Jun03:*

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	129	100.0	129	19	AAW69225
2	95	73.6	113	20	AAW76629
3	90	69.8	117	23	ABP42296
4	49	38.0	68	22	ABG22982
5	41	31.8	294	22	ABBI2440
6	24	18.6	33	22	AAW73633
7	14	10.9	119	22	AAU33022
8	11	8.5	16	21	AAW53504
9	8	6.2	115	22	AAU33023

10	8	6.2	155	22	ABBI2358	Human bone marrow
11	8	6.2	244	23	ABW48985	Listeria monocytog
12	8	6.2	328	22	AAW71925	Human olfactory re
13	8	6.2	526	23	AAW47786	Protein encoded by
14	8	6.2	1512	22	AAW78732	Human protein SQ
15	8	6.2	1512	22	AAU04349	Mammalian toxicolo
16	8	6.2	1550	22	AAW79716	Human protein SQ
17	7	5.4	80	21	AAW63145	Human secreted pro
18	7	5.4	213	24	ABP57640	S. murrayanaensis A
19	7	5.4	247	21	AAW43814	Human cancer assoc
20	7	5.4	305	24	ABP75738	Human secretory po
21	7	5.4	366	22	AAW95617	Human protein sequ
22	7	5.4	366	24	ABG76060	Caspase recruitmen
23	7	5.4	403	23	ABG93239	C. albicans BAX-as
24	7	5.4	424	23	ABP63135	lovf gene expressi
25	7	5.4	424	23	ABP35719	Fungal ZBC protein
26	7	5.4	449	19	AAW30559	Aspergillus oryzae
27	7	5.4	527	22	AAU04873	Micromonospora eve
28	7	5.4	536	22	AAU01205	Human caspase recr
29	7	5.4	536	23	ABG79553	Human CARD-9. Hom
30	7	5.4	568	22	ABG01439	Novel human diagno
31	7	5.4	636	19	AAW30557	Aspergillus oryzae
32	7	5.4	636	19	AAW41509	Aspergillus oryzae
33	7	5.4	636	19	AAW41498	5-aminolevulinic a
34	7	5.4	851	22	ABW62061	Drosophila melanog
35	7	5.4	875	22	ABW67844	Drosophila melanog
36	7	5.4	928	21	AAW77292	Streptomyces antib
37	7	5.4	928	21	AAW78844	KSq-ATQ loading di
38	7	5.4	930	23	AAU76934	Hybrid polyketide
39	7	5.4	972	22	ABW96612	Putative P. abyssi
40	7	5.4	1065	22	ABW25035	Novel human diagno
41	7	5.4	1130	22	ABG02249	Novel human diagno
42	7	5.4	2271	22	ABW65616	Drosophila melanog
43	7	5.4	2291	23	AAU75351	Drosophila Rotkehl
44	7	5.4	3167	24	ABW57680	Saccharopolyspora
45	7	5.4	3170	20	AAW39299	SpnC a polyketide
46	7	5.4	3170	22	AAW70967	S. spinosa protein
47	7	5.4	4150	21	AAW92707	S. antibioticus 8,
48	6	4.7	9	22	AAW99097	Vaccine related MH
49	6	4.7	13	16	AAW21312	Glucagon precursor
50	6	4.7	15	15	AAW45734	Interleukin-6 mute
51	6	4.7	15	22	ABW67796	Human ATP-depend
52	6	4.7	16	20	AAW95534	B38 peptide fragm
53	6	4.7	16	21	AAW70048	A. halophila SDMT
54	6	4.7	25	23	AAU90630	Insulin/insulin-li
55	6	4.7	29	23	ABP42107	Human ovarian anti
56	6	4.7	36	17	AAW95643	Oxyntomodulin vari
57	6	4.7	40	21	AAW58531	Lung cancer associ
58	6	4.7	43	22	AAU04106	HD-Zip subfamily I
59	6	4.7	43	22	AAU04107	HD-Zip subfamily I
60	6	4.7	43	22	AAU04146	HD-Zip subfamily I
61	6	4.7	50	22	AAU55242	Propionibacterium
62	6	4.7	56	21	AAW32615	Eucalyptus grandis
63	6	4.7	56	22	AAU44405	Propionibacterium
64	6	4.7	56	23	ABP09534	Human ORX protein
65	6	4.7	57	22	AAU50998	Propionibacterium
66	6	4.7	58	22	AAW76481	Human colon cancer
67	6	4.7	58	22	AAW77202	Human colon cancer
68	6	4.7	59	21	AAW53282	Human colon cancer
69	6	4.7	59	22	AAW74467	Human colon cancer
70	6	4.7	61	22	ABG01258	Novel human diagno
71	6	4.7	61	22	ABG30204	Novel human diagno
72	6	4.7	62	22	AAO04650	Human polypeptide
73	6	4.7	65	21	AAW27598	Arabidopsis thalia
74	6	4.7	68	22	ABW59574	Human testicular a
75	6	4.7	68	22	AAW95277	Human reproductive

ALIGNMENTS

RESULT 1
AAW69225

ID AAW69225 standard; Protein; 129 AA.
 XX AAW69225;
 AC
 XX
 DT 18-FEB-1999 (first entry)
 XX
 DE NADH dehydrogenase subunit NDS-2.
 XX
 KW NADH dehydrogenase subunit; sympathetic nervous system disease; myopathy;
 KW nicotinamide-adenine dinucleotide dehydrogenase; cancer; leukaemia;
 KW immune system disorder; neurodegenerative disease; therapy; NDS-2.
 XX
 OS Homo sapiens.
 XX
 XX WO9831815-A2.
 PN
 XX
 XX 23-JUL-1998.
 PD
 XX
 PF 17-DEC-1997; 97WO-US23970.
 XX
 XX 17-JAN-1997; 97US-0785065.
 PR
 XX
 XX (INCY-) INCYTE PHARM INC.
 PA
 XX
 PI Bandman O, Goli SK, Hillman JL;
 XX
 XX WPI; 1998-414112/35.
 DR
 DR N-PSDB; AAW44787.
 XX
 XX Human nicotinamide-adenine dinucleotide dehydrogenase sub:units -
 PT useful for, e.g. diagnosis, treatment and prevention of cancer,
 PT myopathy, immune system disease and neurodegeneration
 XX
 XX Claim 19; Fig 2; 80pp; English.
 XX
 CC This sequence represents the NADH (reduced nicotinamide-adenine
 CC dinucleotide) dehydrogenase subunit, NDS-2, of the invention. Cells
 CC containing the DNA are used to produce the recombinant subunits.
 CC Antagonists of NDS-1 (typically antisense sequences or ribozymes) are
 CC used to treat or prevent cancer (leukaemia and solid cancers) and immune
 CC system disorders (e.g. asthma, diabetes, rheumatoid arthritis,
 CC osteoporosis and many others); NDS-2 and NDS-4 are used to treat myopathy
 CC (e.g. ophthalmoplegia, myoclonic epilepsy and lactic acidosis), while
 CC their antagonists are used to treat cancer and disease of the sympathetic
 CC nervous system (e.g. hypertension, arrhythmia and migraine); NDS-3 is
 CC also used to treat myopathy and its antagonists to treat cancer and
 CC neurodegenerative disease (e.g. Alzheimer's, Huntington's and Parkinson's
 CC diseases, epilepsy and Down's syndrome). In all cases NDS or their
 CC antagonists may be expressed from gene therapy vectors. AB may be used
 CC therapeutically as antagonist; as immunoassay reagent for diagnosis or
 CC monitoring such diseases; in competitive screening assays for agents that
 CC bind specifically to the subunits, and for affinity purification of the
 CC subunits from natural sources. The DNAs are useful as primers and probes
 CC for diagnosis and monitoring (including detecting predisposition to
 CC cancer); for gene mapping or identifying related sequences, while the
 CC subunits are also used to raise antibodies and to screen for specific
 CC binding agents.
 XX
 SQ Sequence 129 AA;
 Query Match 100.0%; Score 129; DB 19; Length 129;
 Best Local Similarity 100.0%; Pred. No. 1.2e-129;
 Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSFPKYPSSLRLTPETLDPAEYNISPETRAQAELAIRAQLKREYLLQYNDPNRGLI 60
 DB 1 MSFPKYPSSLRLTPETLDPAEYNISPETRAQAELAIRAQLKREYLLQYNDPNRGLI 60
 QY 61 ENPALLRWAYARTINVTNPRPTPKNSLMGALCGFGLPIFYIYIITERDRKELIOEGK 120
 DB 61 ENPALLRWAYARTINVTNPRPTPKNSLMGALCGFGLPIFYIYIITERDRKELIOEGK 120
 QY 121 LDRTHLSY 129

DB 121 LDRTHLSY 129
 RESULT 2
 AAY76629
 ID AAY76629 standard; Protein; 113 AA.
 XX
 AC AAY76629;
 XX
 DT 10-APR-2000 (first entry)
 XX
 DE Human ovarian tumor EST fragment encoded protein 125.
 XX
 XX Expressed sequence tag; EST; human; ovarian tumor; anticancer;
 KW gene therapy; treatment.
 KW
 XX Homo sapiens.
 OS
 XX DE19817557-A1.
 PN
 XX 21-OCT-1999.
 PD
 XX
 PF 09-APR-1998; 98DE-1017557.
 XX
 XX 09-APR-1998; 98DE-1017557.
 PR
 XX (META-) METAGEN GES GENOMFORSCHUNG MBH.
 PA
 XX Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pillarsky C, Dahl E;
 PI
 XX WPI; 1999-591920/51.
 DR
 DR N-PSDB; AAZ77502.
 XX
 XX New nucleic acid sequences expressed in ovarian, and some other, cancer
 PT tissues, and derived polypeptides, for treatment, of ovarian cancer and
 PT identification of therapeutic agents
 XX
 XX Claim 25; Page 295; 310pp; German.
 XX
 CC This invention describes novel nucleic acid (cDNA) sequences (A) which
 CC have anticancer activity and are highly expressed in ovarian tumor
 CC tissue (and some also in testis and breast cancer tissue). The products
 CC of the invention can be used for gene therapy. (A) are used (i) for
 CC recombinant expression of polypeptides (B) and (ii) to isolate complete
 CC genes. (B) are used (i) to identify agents suitable for treatment of
 CC ovarian cancer; (ii) directly for treating this form of cancer
 CC (including expression from gene therapy vectors) and (iii) for generation
 CC of specific antibodies. (A) are identified by assembling ESTs (expressed
 CC sequence tags) from a particular tissue type before comparison of
 CC expression patterns. This allows a significantly longer fragment of the
 CC gene to be revealed, so should reduce the number of failures associated
 CC with the fact that ESTs from different libraries may represent different
 CC parts of the same unknown gene, distorting the estimated frequency of
 CC occurrence in a particular tissue. AAY76505-Y76638 represent protein
 CC fragments encoded by the human ovarian tumor cDNA library derived EST
 CC fragments represented in AAZ77450-277572.
 XX
 SQ Sequence 113 AA;
 Query Match 73.6%; Score 95; DB 20; Length 113;
 Best Local Similarity 100.0%; Pred. No. 2.3e-93;
 Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 35 ERLAIRAQLKREYLLQYNDPNRGLIENPALLRWAYARTINVTNPRPTPKNSLMGALCG 94
 DB 19 ERLAIRAQLKREYLLQYNDPNRGLIENPALLRWAYARTINVTNPRPTPKNSLMGALCG 78
 QY 95 FGLPIFYIYIITERDRKELIOEGKLDRTFHSY 129
 DB 79 FGLPIFYIYIITERDRKELIOEGKLDRTFHSY 113

CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 68 AA;
 Query Match 38.0%; Score 49; DB 22; Length 68;
 Best Local Similarity 100.0%; Pred. No. 2.1e-44;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 61 ENPALLWAYARTINYPNFRTPKNSLGMALCGFGLPIFYIYIKTER 109
 DB 9 ENPALLWAYARTINYPNFRTPKNSLGMALCGFGLPIFYIYIKTER 57

RESULT 5
 ABB12440
 ID ABB12440 standard; Protein; 294 AA.
 XX AC ABB12440;
 XX DT 15-JAN-2002 (first entry)
 XX DE Human bone marrow expressed protein SEQ ID NO: 279.
 XX KW Human; bone marrow; cytostatic; antirheumatic; antiarthritic; vulnary;
 KW antinflammatory; antibacterial; immunosuppressive; vasotropic; cancer;
 KW antiparkinsonian; neuroprotective; nootropic; haemostatic; osteopathic;
 KW antileuc; fungicide; antidiabetic; antisthmatic; antiallergic;
 KW immunostimulant; analgesic; cerebroprotective; antianaemic; infection;
 KW nervous system disorder; autoimmune disorder; inflammation; allergy.
 XX OS Homo sapiens.
 XX PN WO200174836-A1.
 XX PD 11-OCT-2001.
 XX PF 30-MAR-2001; 2001WO-US10472.
 XX PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 PR 23-AUG-2000; 2000US-0649267.
 PR 30-NOV-2000; 2000US-250583P.
 XX (HYSE-) HYSEQ INC.
 PA Tang YT, Liu C, Drmanac RT, Ford JE, Boyle BJ;
 PI WPI; 2001-626375/72.
 DR New bone marrow-expressed nucleic acids and polypeptides, useful for
 XX diagnosis, treatment of inflammatory, autoimmune, neurological, cancer
 PT and increasing hematopoiesis, stem cell survival and bone growth and
 PT remodeling -
 XX Claim 10; Page 331; 380pp; English.

XX The present invention relates to bone marrow expressed polynucleotides
 CC and proteins. These sequences can be used in the treatment of
 CC inflammatory conditions (eg arthritis, Crohn's disease), cancer, central
 CC and peripheral nervous system diseases and neuropathies, such as
 CC Alzheimer's, Parkinson's and Huntington's diseases, spinal cord
 CC disorders, head trauma, cerebrovascular diseases, myeloid and lymphoid
 CC cell disorders, platelet disorders, stem cell disorders, bone
 CC degenerative disorders, autoimmune disorders, for example multiple
 CC sclerosis, diabetes and arthritis, viral and bacterial infections,
 CC allergies, and blood coagulation disorders. The present sequence is a
 CC protein of the invention.

SQ Sequence 294 AA;
 Query Match 31.8%; Score 41; DB 22; Length 294;
 DB 10 KTERDRKEKLIQEGKLDRTFHLISY 33

Best Local Similarity 100.0%; Pred. No. 2.8e-35;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 40 RAQLKREYLLQYNDNRRGLIENPALLRWAYARTINYPNF 80
 DB 119 RAQLKREYLLQYNDNRRGLIENPALLRWAYARTINYPNF 159

RESULT 6
 AAG73633
 ID AAG73633 standard; Protein; 33 AA.
 XX AC AAG73633;
 XX DT 03-SEP-2001 (first entry)
 XX DE Human colon cancer antigen protein SEQ ID NO:4397.
 XX KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
 KW colorectal carcinoma.
 XX OS Homo sapiens.
 XX PN WO200122920-A2.
 XX PD 05-APR-2001.
 XX PF 28-SEP-2000; 2000WO-US26524.
 XX PR 29-SEP-1999; 99US-0157137.
 PR 03-NOV-1999; 99US-0163280.
 XX (HUMA-) HUMAN GENOME SCI INC.

XX Ruben SM, Barash SC, Birse CE, Rosen CA;
 PI WPI; 2001-235357/24.
 DR N-PSDB; AAH33064.
 XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
 PT useful for preventing, diagnosing and/or treating colorectal cancers -
 XX Claim 11; Page 6222-6223; 9803pp; English.

XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
 CC cancer-associated nucleic acid molecules (N) and proteins (P), where
 CC the proteins are collectively known as colon cancer antigens. The colon
 CC cancer antigens have cytostatic activity and can be used in gene
 CC therapy and vaccine production. N and P may be used in the prevention,
 CC diagnosis and treatment of diseases associated with inappropriate P
 CC expression. For example, N and P may be used to treat disorders
 CC associated with decreased expression by rectifying mutations or deletions
 CC in a patient's genome that affect the activity of P by expressing
 CC inactive proteins or to supplement the patients own production of P.
 CC Additionally, N may be used to produce the colon cancer-associated Ps,
 CC by inserting the nucleic acids into a host cell and culturing the cell
 CC to express the proteins. N and P can be used in the prevention, diagnosis
 CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
 CC and AAB77789 represent sequences used in the exemplification of the
 CC present invention.
 CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
 CC missing at time of publication, meaning no sequences are present for
 CC SEQ ID NO:1027 to 1052, 7921 and 7922.

XX Sequence 33 AA;
 Query Match 18.6%; Score 24; DB 22; Length 33;
 Best Local Similarity 100.0%; Pred. No. 5.8e-18;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 106 KTERDRKEKLIQEGKLDRTFHLISY 129
 DB 10 KTERDRKEKLIQEGKLDRTFHLISY 33

RESULT 7
AAU33022
ID AAU33022 standard; Protein; 119 AA.
XX
AC AAU33022;
XX
DT 18-DEC-2001 (first entry)
XX
DE Novel human secreted protein #3513.
XX
KW Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
OS Homo sapiens.
XX
PN WO200179449-A2.
XX
PD 25-OCT-2001.
XX
PF 16-APR-2001; 2001WO-US08656.
XX
PR 18-APR-2000; 2000US-0552929.
XX
PR 26-JAN-2001; 2001US-0770160.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
XX WPI; 2001-611725/70.
XX
PT Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy -
XX
PS Claim 20; Page 699; 765pp; English.
XX
CC The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising
CC the nucleic acids encoding the polypeptides and cells genetically
CC engineered to express them are also useful for producing the proteins.
CC The proteins are useful in genetic vaccination, testing and
CC therapy, and can be used as nutritional supplements. They may be used to
CC increase stem cell proliferation; to regulate haematopoiesis; and in
CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
CC immune suppression and/or stimulation; as anti-inflammatory agents; and
CC in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid
CC sequences of novel human secreted proteins of the invention.
XX
SQ Sequence 119 AA;
Query Match 10.9%; Score 14; DB 22; Length 119;
Best Local Similarity 100.0%; Pred. No. 8.8e-07;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 75 NVYENFRPTPKNSL 88
DB 71 NVYENFRPTPKNSL 84
RESULT 8
AAB53504
ID AAB53504 standard; Protein; 16 AA.
XX
AC AAB53504;
XX

09-MAR-2001 (first entry)
Human colon cancer antigen protein sequence SEQ ID NO:1044.
Human; colon cancer; colon cancer antigen; diagnosis; detection;
identification; cytostatic; cardioactive; neuroprotective; vulnery;
immunomodulatory; muscular; gynaecological; gastrointestinal;
nephrotropic; antiinfective; antibacterial; gene therapy; wound;
neural disorder; immune system disorder; muscular disorder;
reproductive disorder; gastrointestinal disorder; renal disorder;
infectious disease; cardiovascular disorder.
OS Homo sapiens.
XX
XX WO200055351-A1.
XX
XX 21-SEP-2000.
XX
XX 08-MAR-2000; 2000WO-US05883.
XX
XX 12-MAR-1999; 99US-0124270.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM;
XX
XX WPI; 2000-587534/55.
XX
XX N-PSDB; AAC98261.
XX
XX Colon cancer associated gene sequences, referred to as colon cancer
XX antigens, useful for the treatment, prevention, and diagnosis of colon
XX disorders such as colon cancer -
XX
XX Claim 11; Page 1624; 2104pp; English.
XX
XX AAC97991 to AAC98763 encode the human colon cancer associated proteins,
XX called human colon cancer antigens, given in AAB53234 to AAB54006. The
XX human colon cancer antigens can have cytostatic, cardioactive, muscular;
XX neuroprotective, immunomodulatory, gynaecological, gastrointestinal,
XX vulnery, nephrotropic, antiinfective and antibacterial activities, and
XX can be used in gene therapy. The colon cancer antigen polynucleotides,
XX proteins and antibodies to the proteins are useful for the prevention,
XX treatment and diagnosis of colon disorders, such as colon cancer. The
XX polynucleotides may be used in diagnostics and research, such as for
XX chromosome identification, and as hybridisation probes. The proteins
XX may also be used to prevent diseases such as neural disorders, immune
XX system disorders, muscular disorders, reproductive disorders,
XX gastrointestinal disorders, wounds, renal disorders, infectious
XX diseases, and cardiovascular disorders. AAC98764 to AAC98772 and
XX AAB54007 represent sequences used in the exemplification of the present
XX invention.
SQ Sequence 16 AA;
Query Match 8.5%; Score 11; DB 21; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.00023;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 119 GKLDRTFHLIS 129
DB 6 GKLDRTFHLIS 16
RESULT 9
AAU33023
ID AAU33023 standard; Protein; 115 AA.
XX
XX AAU33023;
XX
XX 18-DEC-2001 (first entry)
XX
XX Novel human secreted protein #3514.
XX

KW Human; vaccination; gene therapy; nutritional supplement;
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
 XX Homo sapiens.
 XX WO200179449-A2.
 XX 25-OCT-2001.
 XX 16-APR-2001; 2001WO-US08656.
 XX 18-APR-2000; 2000US-0552929.
 XX 26-JAN-2001; 2001US-0770160.
 XX (HYSE-) HYSEQ INC.
 XX Tang YT, Liu C, Drmanac RT;
 XX WPI; 2001-611725/70.
 XX Nucleic acids encoding a range of human polypeptides, useful in genetic
 XX vaccination, testing and therapy -
 XX Claim 20; Page 699; 765pp; English.
 XX The invention relates to novel human secreted polypeptides. The
 XX polypeptides and antibodies to the polypeptides are useful for
 XX determining the presence of or predisposition to a disease associated
 XX with altered levels of polypeptide. The polypeptides are also useful for
 XX identifying agents (agonists and antagonists) that bind to them. Cells
 XX expressing the proteins are useful for identifying a therapeutic agent
 XX for use in treatment of a pathology related to aberrant expression or
 XX physiological interactions of the polypeptide. Vectors comprising
 XX the nucleic acids encoding the polypeptides and cells genetically
 XX engineered to express them are also useful for producing the proteins.
 XX The proteins are useful in genetic vaccination, testing and
 XX therapy, and can be used as nutritional supplements. They may be used to
 XX increase stem cell proliferation; to regulate haematopoiesis; and in
 XX bone, cartilage, tendon and/or nerve tissue growth or regeneration;
 XX immune suppression and/or stimulation; as anti-inflammatory agents; and
 XX in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid
 XX sequences of novel human secreted proteins of the invention.
 XX SQ Sequence 115 AA;
 Query Match 6.2%; Score 8; DB 22; Length 115;
 Best Local Similarity 100.0%; Pred. No. 2.2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 75 NVYPNFRP 82
 DB 67 NVYPNFRP 74
 RESULT 10
 ABB12358
 ID ABB12358 standard; Protein; 155 AA.
 AC ABB12358;
 XX 15-JAN-2002 (first entry)
 XX Human bone marrow expressed protein SEQ ID NO: 112.
 XX Human; bone marrow; cytostatic; antirheumatic; antiarthritic; vulnary;
 KW antiinflammatory; antibacterial; immunosuppressive; vasotropic; cancer;
 KW antiparkinsonian; neuroprotective; nootropic; haemostatic; osteopathic;
 KW antitumor; fungicide; antidiabetic; antisthmatic; antiallergic;
 KW immunostimulant; analgesic; cerebroprotective; antianaemic; infection;
 KW nervous system disorder; autoimmune disorder; inflammation; allergy.
 XX Homo sapiens. OS

XX WO200174836-A1.
 XX 11-OCT-2001.
 XX 30-MAR-2001; 2001WO-US10472.
 XX 31-MAR-2000; 2000US-0540217.
 XX 23-AUG-2000; 2000US-0649167.
 XX 23-AUG-2000; 2000US-0649267.
 XX 30-NOV-2000; 2000US-250583P.
 XX (HYSE-) HYSEQ INC.
 XX Tang YT, Liu C, Drmanac RT, Ford JE, Boyle BJ;
 XX WPI; 2001-626375/72.
 XX N-PSDB; ABA09602.
 XX New bone marrow-expressed nucleic acids and polypeptides, useful for
 XX diagnosis, treatment of inflammatory, autoimmune, neurological, cancer
 XX and increasing hematopoiesis, stem cell survival and bone growth and
 XX remodeling -
 XX Claim 10; Page 221; 380pp; English.
 XX The present invention relates to bone marrow expressed polynucleotides
 XX and proteins. These sequences can be used in the treatment of
 XX inflammatory conditions (eg arthritis, Crohn's disease), cancer, central
 XX and peripheral nervous system diseases and neuropathies, such as
 XX Alzheimer's, Parkinson's and Huntington's diseases, spinal cord
 XX disorders, head trauma, cerebrovascular diseases, myeloid and lymphoid
 XX cell disorders, platelet disorders, stem cell disorders, bone
 XX degenerative disorders, autoimmune disorders, for example multiple
 XX sclerosis, diabetes and arthritis, viral and bacterial infections,
 XX allergies and blood coagulation disorders. The present sequence is a
 XX protein of the invention.
 XX SQ Sequence 155 AA;
 Query Match 6.2%; Score 8; DB 22; Length 155;
 Best Local Similarity 100.0%; Pred. No. 2.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSFPKYPK 8
 DB 80 MSFPKYPK 87
 RESULT 11
 ABB48985
 ID ABB48985 standard; Protein; 244 AA.
 XX ABB48985;
 XX 05-FEB-2002 (first entry)
 XX Listeria monocytogenes protein #1689.
 XX Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
 KW vitamin B12; bacterial infection; disease.
 KW Listeria monocytogenes.
 XX WO200177335-A2.
 XX 18-OCT-2001.
 XX 11-APR-2001; 2001WO-FR01118.
 XX 11-APR-2000; 2000FR-0004629.
 XX (INSP) INST PASTEUR. PA

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 09:17:12 ; Search time 33 Seconds

(without alignments)
812.403 Million cell updates/sec

Title: US-09-726-899-3

Perfect score: 129

Sequence: 1 MSFPKYPSSRLTLPETLDP.....DRKELIQEKLDRTHLSY 129

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 789580 seqs, 207824079 residues

Word size : 0

Total number of hits satisfying chosen parameters: 789580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 75 summaries

Database : Published Applications AA.*

- 1: /cgm2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgm2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 3: /cgm2_6/ptodata/1/pubpaa/US05_NEW_PUB.pep.*
- 4: /cgm2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgm2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgm2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 7: /cgm2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgm2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgm2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgm2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgm2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgm2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgm2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgm2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgm2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgm2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgm2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgm2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	129	100.0	129	9	US-09-726-899-3
2	90	69.8	117	12	US-10-264-049-3428
3	24	18.6	33	15	US-10-106-698-4407
4	17	13.2	129	9	US-09-726-899-10
5	11	8.5	16	9	US-09-925-299-1044
6	11	8.5	16	11	US-09-925-299-1044
7	9	7.0	479	15	US-10-156-761-13444
8	6.2	512	12	12	US-09-834-434-2
9	8	6.2	8026	12	US-10-132-134-12
10	7	5.4	143	14	US-10-032-159A-6
11	7	5.4	213	15	US-10-187-267A-65
12	7	5.4	247	9	US-09-925-301-1259
13	7	5.4	366	12	US-10-369-493-17015
14	7	5.4	366	14	US-10-032-159A-2
15	7	5.4	366	14	US-10-032-159A-20

ALIGNMENTS

RESULT 1
US-09-726-899-3
; Sequence 3, Application US/09726899
; Patent No. US20010041356A1
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goli, Surya K.
; APPLICANT: Hillman, Jennifer L.

Sequence 104, Appl
Sequence 9715, Ap
Sequence 7566, Ap
Sequence 4807, Ap
Sequence 5, Appl
Sequence 2683, Ap
Sequence 3, Appl
Sequence 7964, Ap
Sequence 2, Appl
Sequence 7962, Ap
Sequence 7965, Ap
Sequence 200, App
Sequence 1098, Ap
Sequence 1098, Ap
Sequence 1098, Ap
Sequence 3239, Ap
Sequence 869, App
Sequence 31113, A
Sequence 33241, A
Sequence 7255, Ap
Sequence 7978, Ap
Sequence 822, App
Sequence 822, App
Sequence 5241, Ap
Sequence 78, Appl
Sequence 85, Appl
Sequence 92, Appl
Sequence 3935, Ap
Sequence 1919, Ap
Sequence 1919, Ap
Sequence 1919, Ap
Sequence 1919, Ap
Sequence 6134, Ap
Sequence 219, App
Sequence 21, Appl
Sequence 30374, A
Sequence 3075, Ap
Sequence 1047, Ap
Sequence 1047, Ap
Sequence 221, App
Sequence 424, App
Sequence 3466, A
Sequence 2479, Ap
Sequence 3070, Ap
Sequence 259, App
Sequence 259, App
Sequence 6664, Ap
Sequence 6664, Ap
Sequence 48207, A
Sequence 951, App
Sequence 464, App
Sequence 313, App
Sequence 4667, Ap
Sequence 568, App
Sequence 191, App
Sequence 11810, A
Sequence 3502, Ap
Sequence 48, Appl
Sequence 29046, A
Sequence 480, App
Sequence 31964, A

US-10-029-180-104
US-10-369-493-9715
US-10-369-493-7566
US-10-369-493-4807
US-09-798-412-5
US-10-325-917-5
US-10-369-493-2683
US-10-607-809-3
US-10-156-761-7964
US-09-808-880-2
US-10-156-761-7962
US-10-156-761-7965
US-10-239-313A-200
US-09-962-756-1098
US-10-253-471-1098
US-10-264-049-3239
US-09-925-302-869
US-10-029-386-31113
US-10-029-386-32241
US-10-106-698-7578
US-09-925-299-822
US-09-925-299-822
US-10-106-698-5241
US-10-435-766-78
US-10-435-766-85
US-09-764-891-3935
US-09-796-692-1919
US-10-057-475B-1919
US-10-154-884B-1919
US-10-040-862-1919
US-10-106-698-6134
US-10-195-730-219
US-10-289-762-21
US-10-029-386-30374
US-10-264-049-3075
US-09-925-299-1047
US-09-925-299-1047
US-10-195-730-221
US-09-864-408A-424
US-10-104-047-2479
US-09-864-761-33466
US-10-264-049-3070
US-09-764-687-259
US-10-073-961-259
US-09-864-408A-6664
US-09-864-761-48207
US-10-289-762-951
US-10-102-806-464
US-10-114-893-313
US-10-108-260A-4667
US-09-867-550-568
US-09-789-561-191
US-10-369-493-11810
US-10-108-260A-3502
US-09-861-451A-48
US-10-029-386-29046
US-10-103-313-480
US-10-029-386-31964

7 5.4 424 14
7 5.4 443 12
7 5.4 477 12
7 5.4 479 12
7 5.4 536 11
7 5.4 536 15
7 5.4 628 12
7 5.4 929 12
7 5.4 3564 15
7 5.4 4150 11
7 5.4 6145 15
7 5.4 7746 15
9 12
25 12
25 12
29 12
29 12
40 9
50 12
51 12
51 12
58 15
58 15
59 9
59 11
59 15
63 12
63 12
63 12
68 11
78 10
78 12
78 12
78 15
80 15
82 12
82 12
83 12
84 12
87 12
87 12
92 9
92 11
92 11
95 12
101 12
105 12
109 9
110 12
113 9
113 15
115 12
116 9
117 12
121 15
121 15
125 14
126 12
141 9
142 9
142 9
144 12
145 12
146 9
161 12
162 15
163 12

;; TITLE OF INVENTION: NOVEL SUBUNITS OF NADH DEHYDROGENASE
;; NUMBER OF SEQUENCES: 12
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Incyte Pharmaceuticals, Inc.
;; STREET: 3174 Porter Drive
;; CITY: Palo Alto
;; STATE: CA
;; COUNTRY: USA
;; ZIP: 94304
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: DOS
;; SOFTWARE: FastSeq for Windows Version 2.0
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/726,899
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/785,065
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Billings, Lucy J.
;; REGISTRATION NUMBER: 36,749
;; REFERENCE/DOCKET NUMBER: PF-0187 US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 415-855-0555
;; TELEFAX: 415-845-4166
;; TELEX:
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 129 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; IMMEDIATE SOURCE:
;; LIBRARY: Consensus
;; CLONE: Consensus
;; US-09-726-899-3

Query Match 100.0%; Score 129; DB 9; Length 129;
Best Local Similarity 100.0%; Pred. No. 1.3e-119;
Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSFPKPKSLRTPETLDPAEYNISPTERRAQERLAIRAQLKREYLLQYNDPNRGLI 60
DB 1 MSFPKPKSLRTPETLDPAEYNISPTERRAQERLAIRAQLKREYLLQYNDPNRGLI 60

QY 61 ENPALLRWAYARTINYPNFRPTPKNSLMGALCGFGPLFIYIIKTERDRKEKLIQEGK 120
DB 61 ENPALLRWAYARTINYPNFRPTPKNSLMGALCGFGPLFIYIIKTERDRKEKLIQEGK 120

QY 121 LDRTHLSY 129
DB 121 LDRTHLSY 129

RESULT 2
US-10-264-049-3428
; Sequence 3428, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133PI
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1

;; SEQ ID NO 3428
;; LENGTH: 117
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: MISC_FEATURE
;; LOCATION: (3)
;; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
;; FEATURE:
;; NAME/KEY: MISC_FEATURE
;; LOCATION: (27)
;; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
US-10-264-049-3428

Query Match 69.8%; Score 90; DB 12; Length 117;
Best Local Similarity 100.0%; Pred. No. 4.1e-81;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 RTLPETLDPAEYNISPTERRAQERLAIRAQLKREYLLQYNDPNRGLIENPALLRWAYA 71
DB 28 RTLPETLDPAEYNISPTERRAQERLAIRAQLKREYLLQYNDPNRGLIENPALLRWAYA 87

QY 72 RTINYPNFRPTPKNSLMGALCGFGPLFI 101
DB 88 RTINYPNFRPTPKNSLMGALCGFGPLFI 117

RESULT 3
US-10-106-698-4407
; Sequence 4407, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypepti
; FILE REFERENCE: PA005PI
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 4407
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-106-698-4407

Query Match 18.6%; Score 24; DB 15; Length 33;
Best Local Similarity 100.0%; Pred. No. 2.4e-16;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 KTERDRKEKLIQEGKLDRTFHLISY 129
DB 10 KTERDRKEKLIQEGKLDRTFHLISY 33

RESULT 4
US-09-726-899-10
; Sequence 10, Application US/09726899
; Patent No. US20010041356A1
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goli, Surya K.
; APPLICANT: Hillman, Jennifer L.
; TITLE OF INVENTION: NOVEL SUBUNITS OF NADH DEHYDROGENASE
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive

;; CITY: Palo Alto
;; STATE: CA
;; COUNTRY: USA
;; ZIP: 94304
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: DOS
;; SOFTWARE: FastSeq for Windows Version 2.0
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/726,899
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/785,065
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Billings, Lucy J.
;; REGISTRATION NUMBER: 36,749
;; REFERENCE/DOCKET NUMBER: PP-0187 US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 415-855-0555
;; TELEFAX: 415-845-4166
;; TELEX:
;; INFORMATION FOR SEQ ID NO: 10:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 129 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; IMMEDIATE SOURCE:
;; LIBRARY: GenBank
;; CLONE: 114
US-09-726-899-10

Query Match 13.2%; Score 17; DB 9; Length 129;
Best Local Similarity 100.0%; Pred. No. 6.2e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 109 RDRKEKLIQEGKLDRTF 125
Db 109 RDRKEKLIQEGKLDRTF 125

RESULT 5
US-09-925-299-1044
; Sequence 1044, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PAL02
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 1044
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-299-1044

Query Match 8.5%; Score 11; DB 9; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.00094;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 119 GKLDRTFHLISY 129
Db 6 GKLDRTFHLISY 16

RESULT 6
US-09-925-299-1044
; Sequence 1044, Application US/09925299
; Publication No. US20030040617A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PAL02
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 1044
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-299-1044

Query Match 8.5%; Score 11; DB 11; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.00094;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 119 GKLDRTFHLISY 129
Db 6 GKLDRTFHLISY 16

RESULT 7
US-10-156-761-13444
; Sequence 13444, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 13444
; LENGTH: 479
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-13444

Query Match 7.0%; Score 9; DB 15; Length 479;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 LPETLDPAE 22
Db 242 LPETLDPAE 250

RESULT 8
US-09-834-434-2
; Sequence 2, Application US/09834434
; Publication No. US20030162218A1

GENERAL INFORMATION:
; APPLICANT: EMALFARB, MARK A.
; APPLICANT: PUNT, PETER J.
; APPLICANT: VAN ZEIJL, CORNELIA
; APPLICANT: VAN DEN HONDEL, CORNELIUS
; TITLE OF INVENTION: HIGH-THROUGHPUT SCREENING OF EXPRESSED DNA LIBRARIES IN
; FILE REFERENCE: 3123-4006
; CURRENT APPLICATION NUMBER: US/09/834,434
; PRIOR FILING DATE: 2003-03-13
; PRIOR APPLICATION NUMBER: PCT/US00/10199
; PRIOR FILING DATE: 2000-04-13
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 512
; TYPE: PRT
; ORGANISM: Aspergillus niger
US-09-834-434-2

Query Match 6.2%; Score 8; DB 12; Length 512;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 PSSRLTLP 15
|||
DB 186 PSSRLTLP 193

RESULT 9
US-10-132-134-12
; Sequence 12, Application US/101321134
; Publication No. US20030171562A1
; GENERAL INFORMATION:
; APPLICANT: Farnet, Chris
; APPLICANT: Yang, Xianshu
; APPLICANT: Staffa, Alfredo
; APPLICANT: Zazopoulos, Emmanuel
; TITLE OF INVENTION: POLYKETIDE SYNTHASE ENZYMES
; FILE REFERENCE: 3012-2US
; CURRENT APPLICATION NUMBER: US/10/132,134
; CURRENT FILING DATE: 2002-04-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 8026
; TYPE: PRT
; ORGANISM: Streptomyces platensis subsp. roseaceus
US-10-132-134-12

Query Match 6.2%; Score 8; DB 12; Length 8026;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 RRAQAERL 37
|||
DB 5943 RRAQAERL 5950

RESULT 10
US-10-032-159A-6
; Sequence 6, Application US/10032159A
; Publication No. US20020164703A1
; GENERAL INFORMATION:
; APPLICANT: Pawlowski, Krzysztof
; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; TITLE OF INVENTION: CARD-DOMAIN CONTAINING POLYPEPTIDES,
; FILE REFERENCE: P-LJ 5100
; CURRENT APPLICATION NUMBER: US/10/032,159A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: US 60/257,457

PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-032-159A-6

Query Match 5.4%; Score 7; DB 14; Length 144;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 117 QEGKLDLR 123
|||
DB 116 QEGKLDLR 122

RESULT 11
US-10-187-267A-65
; Sequence 65, Application US/10187267A
; Publication No. US20030124679A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Paradkar, Ashish
; APPLICANT: Varoglu, Mustafa
; APPLICANT: Mathur, Eric J.
; TITLE OF INVENTION: GLYCOSYLATED KINAMYCINS AND METHODS OF
; TITLE OF INVENTION: MAKING AND USING THEM
; FILE REFERENCE: 09010-280001
; CURRENT APPLICATION NUMBER: US/10/187,267A
; CURRENT FILING DATE: 2003-01-27
; PRIOR APPLICATION NUMBER: US 60/301,401
; PRIOR FILING DATE: 2001-06-27
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 65
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Streptomyces murayamaensis ATCC 21414
US-10-187-267A-65

Query Match 5.4%; Score 7; DB 15; Length 213;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 ETLDPAE 22
|||
DB 39 ETLDPAE 45

RESULT 12
US-09-925-301-1259
; Sequence 1259, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1259
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 09:14:57 ; Search time 21 Seconds

(without alignments)
259.910 Million cell updates/sec

Title: US-09-726-899-3

Perfect score: 129

Sequence: 1 MSPFKYKPSRLTLPETLDP.....DRKEKLIQEGKLDRTTHLSY 129

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 328717 seqs, 42310858 residues

Word size : 0

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 75 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/iaa/5A COMB.pep:*
- 2: /cgn2_6/ptodata/1/iaa/5B COMB.pep:*
- 3: /cgn2_6/ptodata/1/iaa/6A COMB.pep:*
- 4: /cgn2_6/ptodata/1/iaa/6B COMB.pep:*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pep:*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	ID	Description
1	129	100.0	129	2	US-08-785-065-3
2	129	100.0	129	4	US-09-151-412-3
3	17	13.2	129	2	US-08-785-065-10
4	17	13.2	129	4	US-09-151-412-10
5	8	6.2	1112	4	US-09-252-991A-27256
6	8	6.2	1440	3	US-09-357-251-37
7	8	6.2	1512	4	US-09-443-184-48
8	7	5.4	301	4	US-09-252-991A-28663
9	7	5.4	449	2	US-08-819-458A-16
10	7	5.4	455	4	US-09-252-991A-22154
11	7	5.4	636	2	US-08-871-266B-2
12	7	5.4	636	2	US-08-819-458A-2
13	7	5.4	636	2	US-09-018-864A-2
14	7	5.4	636	3	US-08-871-267B-2
15	7	5.4	636	3	US-09-618-419-2
16	7	5.4	3170	3	US-09-036-987A-4
17	7	5.4	3170	3	US-09-370-700-4
18	7	5.4	3170	4	US-09-603-207-4
19	7	5.4	4150	3	US-09-428-517-2
20	6	4.7	15	1	US-07-918-181A-30
21	6	4.7	15	1	US-08-231-575-30
22	6	4.7	15	5	PCT-US93-06928-30
23	6	4.7	16	2	US-08-675-921B-6
24	6	4.7	36	2	US-08-836-528-1
25	6	4.7	83	4	US-09-198-452A-21
26	6	4.7	92	1	US-08-446-038B-21
27	6	4.7	92	1	US-08-446-010B-21

28	6	4.7	92	2	US-08-479-078-18	Sequence 18, Appl
29	6	4.7	92	2	US-08-805-445-21	Sequence 21, Appl
30	6	4.7	92	2	US-08-064-067D-21	Sequence 21, Appl
31	6	4.7	92	2	US-09-066-208-21	Sequence 21, Appl
32	6	4.7	94	1	US-08-167-035-23	Sequence 23, Appl
33	6	4.7	94	1	US-08-208-887A-23	Sequence 23, Appl
34	6	4.7	94	2	US-08-539-005-23	Sequence 23, Appl
35	6	4.7	94	4	US-08-280-598-25	Sequence 25, Appl
36	6	4.7	115	4	US-09-107-532A-4792	Sequence 4792, Ap
37	6	4.7	117	4	US-09-198-452A-951	Sequence 951, App
38	6	4.7	123	4	US-09-732-210-1162	Sequence 1162, Ap
39	6	4.7	144	1	US-07-956-700B-89	Sequence 89, Appl
40	6	4.7	144	1	US-08-476-537-89	Sequence 89, Appl
41	6	4.7	144	1	US-08-485-607-89	Sequence 89, Appl
42	6	4.7	144	2	US-08-475-879-89	Sequence 89, Appl
43	6	4.7	144	4	US-09-433-043B-89	Sequence 89, Appl
44	6	4.7	152	4	US-09-252-991A-25743	Sequence 25743, A
45	6	4.7	154	4	US-09-252-991A-29786	Sequence 29786, A
46	6	4.7	157	3	US-09-461-474-14	Sequence 14, Appl
47	6	4.7	161	1	US-07-991-867B-5	Sequence 5, Appl
48	6	4.7	161	1	US-08-107-755A-5	Sequence 5, Appl
49	6	4.7	161	2	US-08-544-332-5	Sequence 5, Appl
50	6	4.7	161	4	US-09-370-861A-5	Sequence 5, Appl
51	6	4.7	166	4	US-09-252-991A-18131	Sequence 18131, A
52	6	4.7	180	4	US-09-732-210-522	Sequence 522, App
53	6	4.7	195	4	US-09-071-035-336	Sequence 336, App
54	6	4.7	204	4	US-09-107-532A-5769	Sequence 5769, Ap
55	6	4.7	227	4	US-09-323-872A-10	Sequence 10, Appl
56	6	4.7	227	4	US-09-072-433-19	Sequence 19, Appl
57	6	4.7	229	5	PCT-US96-03916-13	Sequence 13, Appl
58	6	4.7	229	5	PCT-US96-03916-62	Sequence 62, Appl
59	6	4.7	239	4	US-09-325-932A-65	Sequence 65, Appl
60	6	4.7	242	4	US-08-220-602B-23	Sequence 23, Appl
61	6	4.7	244	4	US-09-328-352-4316	Sequence 4316, Ap
62	6	4.7	250	4	US-09-252-991A-20223	Sequence 20223, A
63	6	4.7	254	4	US-09-252-991A-23519	Sequence 23519, A
64	6	4.7	260	2	US-08-675-921B-9	Sequence 9, Appl
65	6	4.7	262	4	US-09-252-991A-21023	Sequence 21023, A
66	6	4.7	262	4	US-09-252-991A-27182	Sequence 27182, A
67	6	4.7	270	4	US-09-252-991A-25871	Sequence 25871, A
68	6	4.7	292	4	US-09-328-352-8013	Sequence 8013, Ap
69	6	4.7	293	4	US-09-252-991A-25265	Sequence 25265, A
70	6	4.7	301	4	US-09-215-694-13	Sequence 13, Appl
71	6	4.7	322	4	US-09-252-991A-27538	Sequence 27538, A
72	6	4.7	324	4	US-08-732-749-4	Sequence 4, Appl
73	6	4.7	324	4	US-09-258-613-2	Sequence 2, Appl
74	6	4.7	339	4	US-09-107-532A-4716	Sequence 4716, Ap
75	6	4.7	343	4	US-09-107-532A-4619	Sequence 4619, Ap

ALIGNMENTS

RESULT 1
US-08-785-065-3
; Sequence 3, Application US/08785065
; Patent No. 5814451
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goli, Surya K.
; APPLICANT: Hillman, Jennifer L.
; TITLE OF INVENTION: NOVEL SUBUNITS OF NADH DEHYDROGENASE
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible

```

;
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/785,065
; FILING DATE: Herewith
; CLASSIFICATION: 424
; PRIOR APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0187 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
;
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 129 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: Consensus
; CLONE: Consensus
;
US-09-785-065-3
;
Query Match 100.0%; Score 129; DB 2; Length 129;
Best Local Similarity 100.0%; Pred. No. 5.7e-119;
Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSFPKYKPSLRLTPETLDPAEYNISPETTRQAERLAIRAOIKREYLLQYNDPNRGLI 60
DB 1 MSFPKYKPSLRLTPETLDPAEYNISPETTRQAERLAIRAOIKREYLLQYNDPNRGLI 60

QY 61 ENPALLWAYARTINYPNFRPTPKNSLMGALCGFGLFIYIIKTERDRKEKLIQEGK 120
DB 61 ENPALLWAYARTINYPNFRPTPKNSLMGALCGFGLFIYIIKTERDRKEKLIQEGK 120

QY 121 LDRTHLSY 129
DB 121 LDRTHLSY 129

RESULT 2
US-09-151-412-3
; Sequence 3, Application US/09151412
; Patent No. 6399345
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goli, Surya K.
; APPLICANT: Hillman, Jennifer L.
; TITLE OF INVENTION: NOVEL SUBUNITS OF NADH DEHYDROGENASE
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/151,412
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/785,065

;
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/785,065
; FILING DATE: Herewith
; CLASSIFICATION: 424
; PRIOR APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0187 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
;
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 129 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: Consensus
; CLONE: Consensus
;
US-09-151-412-3
;
Query Match 100.0%; Score 129; DB 4; Length 129;
Best Local Similarity 100.0%; Pred. No. 5.7e-119;
Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSFPKYKPSLRLTPETLDPAEYNISPETTRQAERLAIRAOIKREYLLQYNDPNRGLI 60
DB 1 MSFPKYKPSLRLTPETLDPAEYNISPETTRQAERLAIRAOIKREYLLQYNDPNRGLI 60

QY 61 ENPALLWAYARTINYPNFRPTPKNSLMGALCGFGLFIYIIKTERDRKEKLIQEGK 120
DB 61 ENPALLWAYARTINYPNFRPTPKNSLMGALCGFGLFIYIIKTERDRKEKLIQEGK 120

QY 121 LDRTHLSY 129
DB 121 LDRTHLSY 129

RESULT 3
US-08-785-065-10
; Sequence 10, Application US/08785065
; Patent No. 5814451
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goli, Surya K.
; APPLICANT: Hillman, Jennifer L.
; TITLE OF INVENTION: NOVEL SUBUNITS OF NADH DEHYDROGENASE
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/785,065
; FILING DATE: Herewith
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0187 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166

```